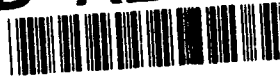


AD-A251 593

PL-TR-91-2163



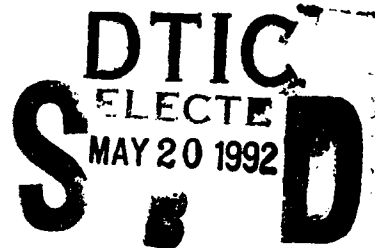
2

PROCESSING OF CLOUD DATABASES FOR THE DEVELOPMENT
OF AN AUTOMATED GLOBAL CLOUD CLIMATOLOGY

James H. Willand and Julia Steeves

ST Systems Corporation
109 Massachusetts Avenue
Lexington, MA 02173

30 June 1991



Scientific Report No. 1

Approved for public release; distribution unlimited.



PHILLIPS LABORATORY
AIR FORCE SYSTEMS COMMAND
HANSCOM AFB, MASSACHUSETTS 01731-5000

92-13364




92 5 19 088

"This technical report has been reviewed and is approved for publication"


DONALD D. GRANTHAM
Contract Manager


DONALD D. GRANTHAM
Branch Chief


ROBERT A. MCCLATCHEY
Division Director

This document has been reviewed by the ESD Public Affairs Office (PA) and is releasable to the National Technical Information Service (NITS).

Qualified requesters may obtain additional copies from the Defense Technical Information Center. All others should apply to the National Technical Information Service.

If your address has changed, or if you wish to be removed from the mailing list, or if the addressee is no longer employed by your organization, please notify PL/IMA, Hanscom AFB, MA 01731-5000. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligation or notices on a specific document requires that it be returned.

DISCLAIMER NOTICE

**THIS DOCUMENT IS BEST QUALITY
PRACTICABLE. THE COPY FURNISHED
TO DTIC CONTAINED A SIGNIFICANT
NUMBER OF PAGES WHICH DO NOT
REPRODUCE LEGIBLY.**

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
<small>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to: Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.</small>				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 30 June 1991	3. REPORT TYPE AND DATES COVERED Scientific Report No. 1	
4. TITLE AND SUBTITLE PROCESSING OF CLOUD DATABASES FOR THE DEVELOPMENT OF AN AUTOMATED GLOBAL CLOUD CLIMATOLOGY			5. FUNDING NUMBERS PE 62101F PR 6670 TA 09 WU AP Contract F19628-88-C-0089	
6. AUTHOR(S) James H. Willand and Julia Steeves				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) ST Systems Corporation 109 Massachusetts Avenue Lexington, MA 02173			8. PERFORMING ORGANIZATION REPORT NUMBER STX Scientific Report No. 1	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Phillips Laboratory Hanscom AFB, MA 01731-5000 Contract Manager: Don Grantham/LYA			10. SPONSORING/MONITORING AGENCY REPORT NUMBER PL-TR-91-2163	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release. Distribution unlimited.			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) This report describes data sources, data processing activities, compression, prediction, and blending of cloud databases being utilized in the development of an automated fast retrieval global cloud climatology. Correlation of sky cover over a sky dome from uplooking whole sky photos and from downviewing satellite data is demonstrated. An updated documentation of a whole sky photo database and documentation of two geographic databases optimized for use on small PCs are also provided. Several FORTRAN subroutines pertinent to cloud data processing are listed in the Appendix.				
14. SUBJECT TERMS Global cloud climatology Simulation Predictions Frequency distributions Data compaction Mean Correlation Sky/cloud cover			15. NUMBER OF PAGES 134	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited Distribution	

ACKNOWLEDGMENTS

The authors would like to acknowledge Mr. Don Grantham, Branch Chief of the LYA section of the Phillips Laboratory, Geophysics Directorate, for contract funding and in-house use of computers and office space. We would like to recognize Mr. A. Boehm, ST Systems Corporation, for his technical assistance and Ms. Norma Tocco, also of ST Systems Corporation, for the technical editing and preparation of this manuscript.



Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

TABLE OF CONTENTS

1.	INTRODUCTION.....	1-1
2.	DOE CLOUD DATABASE DESCRIPTION.....	2-1
2.1	Processing of DOE Cloud Data.....	2-2
2.1.1	Program FETCH.....	2-3
2.1.2	Program MLO.....	2-6
2.1.3	Program W9090.....	2-8
2.1.4	Program MAPDIS.....	2-12
2.1.4.1	Optimum Number of Coefficients for Predicting Parameters.....	2-13
2.1.4.2	Error Analysis of Predicted Mean Sky Cover.....	2-18
2.2	Processing Interannual Variance of Mean Sky Cover.....	2-30
2.2.1	Program FETC2.....	2-30
2.2.2	Program ANULAN,AUNOCE.....	2-32
2.2.3	Program AUN90.....	2-33
3.	NIMBUS 7 SATELLITE CMATRIX CLOUD DATA.....	3-1
3.1	Processing of NIMBUS 7 CMatrix Cloud Data....	3-2
3.1.1	Program EXTRAK.....	3-2
3.1.2	Program SATSUM.....	3-6
3.1.3	Program NORSAT.....	3-6
3.1.4	Program W9090.....	3-7
3.1.5	Program MAPDIS.....	3-8
3.2	Cloud Cover Correlation Structure Using CMatrix Cloud Data.....	3-8
3.3	Processing CMatrix Data for Interannual Variance of Mean Cloud Cover.....	3-12
3.3.1	Program NUSAT.....	3-12
3.3.2	Program A9090.....	3-16
3.3.3	Program MAPDIS.....	3-16
4.	BURGER SKY COVER DATABASE DESCRIPTION.....	4-1
4.1	Processing of Burger Sky Cover Data.....	4-2
4.1.1	Method 1 Program BURFOR.....	4-2
4.1.1.1	Scale Distance to Mean Correlation (Subroutine SKYRO)....	4-2
4.1.2	Method 2 Program BURRHO.....	4-12
4.1.2.1	RUSSWO Distributions to Mean and Variance for New Mean Correlations.....	4-13
4.1.3	Program W9090.....	4-14
4.2	Comparisons of Burger RUSSWO Sky Cover Distributions with CUB predicted Distributions.....	4-16

5.	DATSAV DATABASE DESCRIPTION.....	5-1
5.1	Processing of DATSAV and DATSAV2 Data.....	5-1
5.1.1	Program GETCOV.....	5-2
5.1.2	Formatting for Spatial and Temporal Correlation of Sky Cover.....	5-5
5.1.2.1	Program FORDAT.....	5-5
5.1.3	Formatting DATSAV Data for W9090 Program Runs.....	5-7
5.1.3.1	Program FORDA2.....	5-7
6.	SUMMARY OF CLOUD DATABASES.....	6-1
6.1	Database Blending.....	6-3
6.1.1	Interpolating Data in GMT to Any Time of Day in MAST.....	6-4
6.1.2	Program MAST.....	6-6
6.1.3	Program A9090.....	6-8
6.2	Future Database Blender.....	6-10
7.	WHOLE SKY PHOTOGRAPHS DATA DESCRIPTION.....	7-1
7.1	Whole Sky Photo Data Format.....	7-1
7.2	Processing of Whole Sky Photo Data.....	7-6
7.2.1	WSP Tape Archive Update.....	7-6
7.2.1.1	Program PUTTP.....	7-8
7.2.2	Processing of WSP Data for Sky Dome Cloud Correlation Structure.....	7-9
7.2.2.1	Program CM.....	7-9
7.2.2.2	Program PLOTW.....	7-9
7.2.3	Correlation of Sky Cover Over a Sky Dome.....	7-10
7.2.3.1	Introduction.....	7-10
7.2.3.2	Procedures.....	7-10
7.2.3.3	Results.....	7-11
7.2.3.4	Conclusion.....	7-13
8.	GEOGRAPHY DATABASES.....	8-1
8.1	A Modified NCAR Vector Geography Database for Small PCs.....	8-1
8.2	NAVY Global 10-Minute Raster Elevation Data.....	8-4
8.2.1	Land and Water 10-Minute Data Set.....	8-5
8.2.1.1	Southern Hemisphere.....	8-6
8.2.1.2	Northern Hemisphere.....	8-9
REFERENCES		R-1
APPENDIX A		A-1

1.0 INTRODUCTION

The algorithms for fitting climatological sky cover distributions using only two parameters, mean sky cover and either scale distance or the more recent parameter, mean correlation (see note below), makes possible the development of automated global cloud climatologies. The remaining major requirement for completed automated global cloud climatologies is the processing and/or data compression of massive amounts of global cloud observations from several different cloud data archives for creating a Climatology of Cloud Statistics (C Clouds S). The cloud statistics of interest are mean sky covers and variances, mean correlations of sky cover, sample sizes (populations) and interannual standard deviations of mean sky cover. This document describes the cloud databases being utilized to extract and process the cloud statistics and reports on the progress and data flow of this processing effort.

NOTE: Mean correlation is defined as the mean of the correlations of given conditions between all pairs of points in a domain where a domain can be either a line, an area, or a volume.

The processing activities described in this document start with the data ensemble of DOE/NCAR ground observed and NIMBUS 7 satellite observed mean cloud covers and standard deviations. These two databases have extensively expanded the

original cloud database for automated climatologies to any point on the globe between 90 degrees south to 90 degrees north.

The next processing activities described pertain to the Burger and the DATSAV data ensembles. These cloud databases had to be processed to convert the parameter scale distance to mean correlation. Two methods of conversion are described.

A section is included describing an initial phase of a database blender that may be used to blend all the climatological databases together into one single database for a final automated global cloud climatology.

A description of a whole sky photo database is included mainly to update the documentation of this data set and to show the correlation structure of sky cover over a sky dome which is important information in the future development of sky cover simulators.

Finally, a section on two worldwide geography data sets useful for implementation on small PCs is included for inserting future mapping features such as major land boundaries and terrain induced cloud effects into the automated global cloud climatology.

Several figures showing results of the processing effort are also included.

Most of the software described in this document was written in FORTRAN 5 and 77 for the CDC 6800 Cyber NOS system. Several FORTRAN subroutines and functions critical to the processing effort are listed in Appendix A. A few applications were transferred to a VAX and CONVEX system. These transition points are well documented within the text.

The small PCs that are often referred to are in the computer class of a Zenith 100 or a Zenith 248.

2.0 DOE CLOUD DATABASE DESCRIPTION

The "Climatological Data for Clouds Over the Globe From Surface Observations" (Hahn, 1987) is one of the cloud databases being considered for use in the ongoing development of an automated global cloud climatology. This cloud database was prepared and archived for use in several United States energy agencies including the United States Department of Energy (DOE) and the National Center for Atmospheric Research (NCAR). For convenience, subsequent sections of this report will refer to these data as the DOE cloud database. Hahn's term "cloud cover" is synonymous with "sky cover" as used elsewhere throughout this report. The term "cloud cover" is also used when referring to satellite cloud observed conditions.

The DOE cloud database contains global maps (90°N to 90°S) of long-term monthly and/or seasonal total cloud cover, cloud type amounts and frequencies of occurrence, low cloud base heights, harmonic analysis of annual and diurnal cycles, interannual variations and trends, and cloud type co-occurrences. Most of the data are mapped on grids whereby each grid box represents a 5-degree latitude by 5-degree longitude resolution. Sizes of the grid boxes increase in longitude poleward of 50 degrees latitude which creates nearly equal area grid boxes. According to Hahn (1987), data from all available stations within a grid box were used to compute

the average for that box. The data are archived so that land-observed cloud data are stored separately from ocean-observed data. The period of record for land observations is 11 years from January 1971 through December 1981. The period of record for ship observations is 54 years from January 1930 through November 1983. Portions of the DOE data archive are available in the form of atlases (Warren et al., 1988), and considerable additional information is available on a single magnetic tape.

2.1 Processing of DOE Cloud Data

A single magnetic tape containing the DOE cloud climatology data was acquired by the Geophysics Directorate. ST Systems Corporation was tasked to determine its usefulness in the development of an automated global cloud climatology. A computerized system was implemented to extract selected portions of data from the DOE data tape and then compress the data down to a manageable size for subsequent processing on a PC. The portion of DOE data selected for use in the automated cloud climatology development was the three hourly seasonal mean cloud cover. This mean cloud cover, which was archived eight times in a 24-hour period, has provided valuable information for defining and predicting diurnal cycles of global cloud cover. Processing of the DOE data is described for the winter and summer seasons. (Data for spring and fall seasons are processed operationally using the same procedures.) Seasons mentioned in this report are northern

hemisphere seasons. For example, winter is northern hemisphere winter but southern hemisphere summer.


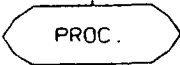
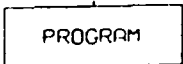

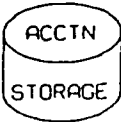

The data processing configuration shown in Figure 2.1 illustrates the sequence of events used to process the DOE winter and summer seasonal mean cloud cover from the initial archive to data compression. This figure and subsequent flow charts demonstrate how data sources are combined with developed software to produce end results. These diagrams all use a common form of symbols defined in Table 2.0. The diagrams show data sources followed by data transfer software developed to format, compress, and transfer data to permanent files. Backup tapes and labels also shown record where all data for each project are kept for permanent retention. Details of the data flow in Figure 2.1 are described below.

2.1.1 Program FETCH

Program FETCH, which is invoked by the procedure XEQFET, extracts global maps of selected cloud statistics from separate files of the DOE cloud climatology tape which have been assigned map group numbers specified by Hahn (1987). Specifically, selected global maps of mean cloud cover for winter at eight GMT times daily over land areas are extracted from file 2 under map group numbers 14 through 21 and summer 30 through 37. Ocean data for the same parameter are extracted from file 6 under map group numbers 1080 through

TABLE 2.0

Symbol Definitions Used in Data Processing
Configuration Diagrams

	MAGNETIC TAPE DATA STORAGE
	PROCEDURE USED FOR BATCH PROCESSING
	COMPUTER PROGRAM (FORTRAN)
	RESULTING PRODUCT OR OTHER PROGRAMS
	ACCOUNT NUMBER PERMANENT DISK DATA STORAGE
	FLOPPY DISK STORAGE

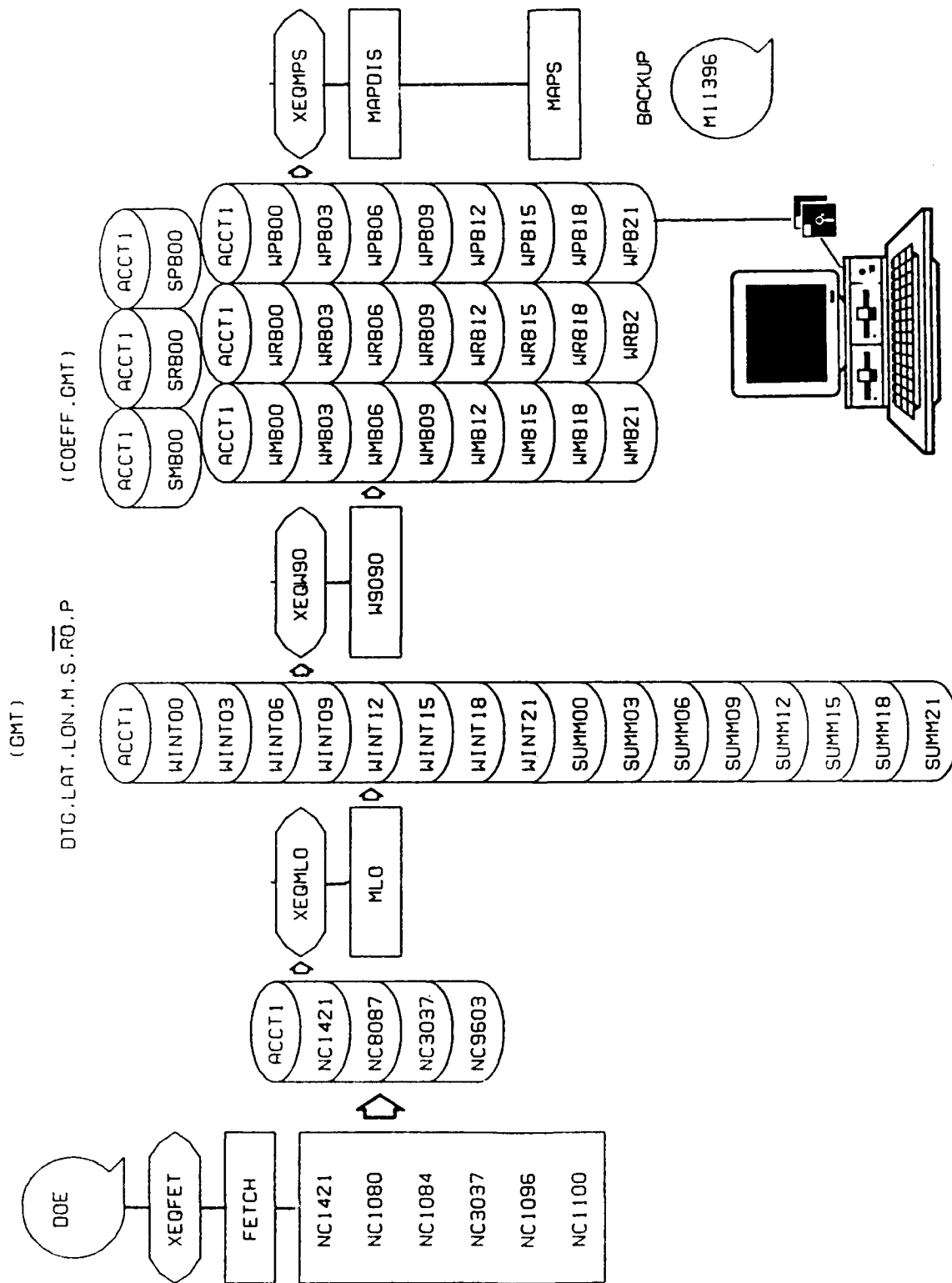


Figure 2.1. Data Flow Configuration Used in Processing the DOE Seasonal Cloud Database.

1087 for winter and 1096 through 1103 for summer. The mean cloud covers for winter over land and ocean areas are stored on permanent disk files under the names NC1421 and NC8087 respectively. The records of mean cloud covers stored for each box also include standard deviations, populations, and central latitude-longitude reference points.

2.1.2 Program MLO

Procedure XEQMLO calls on program MLO for the purpose of merging land with ocean data. During the merging process, the standard deviation of mean cloud cover for each 5-degree x 5-degree grid box is adjusted by

$$SD = SD' \times (N-1)/N \quad (2.1)$$

where

SD' is the DOE standard deviation of mean cloud cover at a grid point

N is the population count (sample size) at a given grid point.

In addition, mean correlation is computed using the given mean cloud cover amount and standard deviation archived for each grid point as input to an accurate tetrachoric correlation function like that listed in Appendix A and described in Smyth (1991). Arguments for this tetrachoric correlation function such as A, B, C, and D, typically represent the normalized

counts in a fourfold contingency table of dichotomous events. But instead, mean correlation (\overline{RO}) can best be computed by setting these arguments to the following computed values.

$$A = V + M \times M \quad (2.2)$$

$$B = M - A \quad (2.3)$$

$$C = B \quad (2.4)$$

$$D = 1.0 - A - 2C \quad (2.5)$$

where

$$M = \text{Mean sky cover} \quad 0 \leq M \leq 1.0$$

$$V = \text{Variance} \quad 0 \leq V \leq .25$$

then

$$\overline{RO} = \text{TETRA}(A, B, C, D)$$

A date time group, latitude-longitude, mean sky cover, standard deviation, mean correlation, and population count for each grid point for each separate season and time period were written to permanent disk storage using the following format.

```
WRITE(NTAPE,100) ITC, ISEA, ITM, ITIM, LAT, LON, M, S, RO, P (2.0)
```

```
100 FORMAT(I3, I2, I1, I2, F6.2, F7.2, 3F6.3, I8)
```

Variables are defined as

ITC: Data type code. DOE data = 4.

ISEA: Season code. 41 = Winter, 43 = Summer.

ITM: Time Code. GMT = 2.
 ITIM: Time. 00,03,06,...21 GMT.
 LAT: Latitude. $-90 \leq \text{LAT} \leq 90$
 LON: Longitude. 0 to 360 degrees East.
 M: Mean cloud cover. $0.0 \leq M \leq 1.0$
 S: Standard deviation. $0.0 \leq S \leq 1.0$
 RO: Mean correlation. $0.0 \leq \overline{R_O} \leq 1.0$
 P: Population

Names such as WINT00, which references data for the winter season at 00 GMT, and SUMM03, which references data for the summer season at 03 GMT, are assigned to the resulting data files. A total of eight files at eight GMTs for each selected season are assembled. The data are now in a form suitable for entry into a data compression program called W9090.

2.1.3 Program W9090

Program W9090 is an upgraded version of program A6060, described by Willand (1988). Both programs enlist a Fourier-Legendre analysis to compress data into the form of spatial coefficients. The Fourier analysis is used longitudinally and the Legendre latitudinally. A discussion of the algorithm can be found in Gerlach (1988).

The W9090 upgrade provides analysis of data between 90 degrees south and 90 north. In addition, a weighting function is introduced to the code forcing values of points having lower population counts to be weighted more heavily toward the values having higher population counts. The coefficients generated by W9090 are subsequently used in the rapid retrieval and prediction of accurate parameter values over any given location of the globe.

Coefficients can be generated by the W9090 program using any reasonable two-dimensional array size. This array size (NF x NL) is set in a parameter statement found at the beginning of the program. NF is the number of desired Fourier terms and NL is the number of Legendre terms. The number of terms determines the number of coefficients that will be generated and the amount of time it will take a Control Data Corporation (CDC) computer running under the NOS system to generate them. The total number of coefficients (LF) generated can be computed by the formula

$$LF = NL \times NF + NL + NF + 1 \quad (2.6)$$

where

NL = Number of Legendre terms

NF = Number of Fourier terms.

The number of seconds required to generate the coefficients given various array sizes is shown in Figure 2.2. Also shown

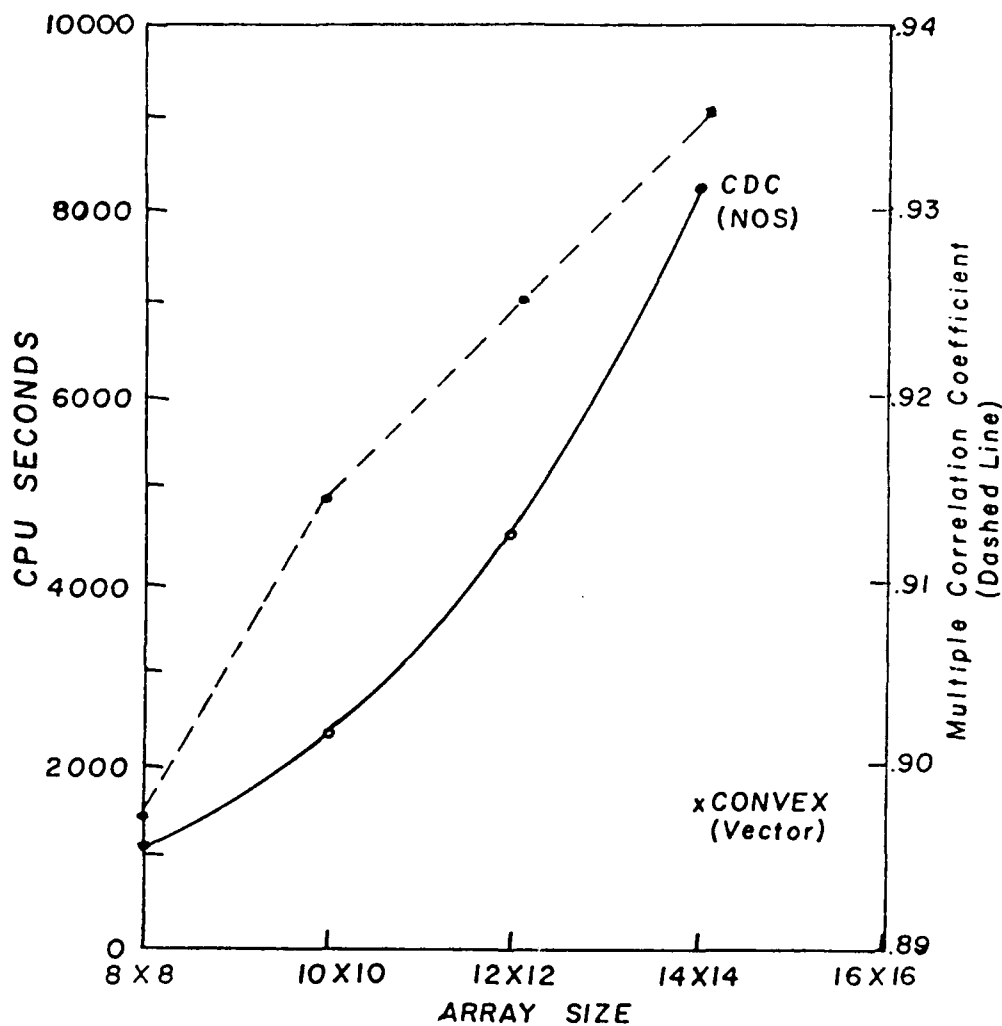


Figure 2.2. Estimated Seconds of Computer Time for Program W9090 to Generate Spatial Coefficients Given Various Fourier-Legendre Array Sizes (solid curve). Improvement of the Multiple Correlation Coefficient Computed by W9090 for Increasing Array Sizes Is Also Shown (dashed curve). The Multiple Correlations Represented Are for DOE Winter Mean Sky Cover at 00 GMT.

is the improvement of the W9090 computed multiple correlation coefficient with increased array size. Considering that many runs will be made using program W9090, the exponential increase in time given increase in array size is of relevant concern. Therefore, a benchmark run was performed using the new CONVEX computer recently installed at the laboratory. This new system uses a vectorization scheme that provides for more efficient matrix computations. As shown in Figure 2.2, the time was considerably reduced when running the 14 x 14 array size. Therefore, future operational W9090 runs will be performed using the CONVEX system.

Coefficients for three parameters are generated. The parameters are 1) mean sky cover, 2) mean correlation, and 3) the natural log of population count. Each set is stored under a different name. Thus, coefficients for the winter (W) season representing mean (M) sky cover from a 14 x 14 (B) array size at 00 GMT (00) are named WMB00. Summer season coefficients are prefixed with an S, $\overline{R0}$ is tagged R, and population P. As mentioned above, the number of coefficients generated depends on the array size chosen. Table 2.1 tabulates the number of coefficients generated by program W9090 given increasing Fourier-Legendre sizes. The FORTRAN format used to store all coefficients is

```
      WRITE (NTAPE,101)I,A(I)
101  FORMAT (I6,E15.7)
```

Variables are

I = Index Counter

A(I) = Coefficient.

2.1.4 Program MAPDIS

Program MAPDIS was designed to read in the coefficients generated by program W9090 and to use these coefficients to predict the parameter of interest: mean sky cover, mean correlation or ln of population over any point on the globe. The predicted values are then processed through a pattern mapping subroutine that uses a bilinear interpolation scheme to generate pattern maps of the predicted parameters.

TABLE 2.1

**Number of Coefficients Generated by W9090 Given
Fourier/Legendre Array Sizes**

Array Size	Number of Coefficients
8 x 8	81
10 x 10	121
12 x 12	169
14 x 14	225
16 x 16	289
18 x 18	361
20 x 20	441

A cylindrical equal area projection of the globe was then overlaid onto the pattern maps for geographic referencing. The equal area projection (48 degrees standard parallel) used resembles a Peters projection (45 degrees standard parallel). The latitude parallels are simply computed as the sine of the latitude. These pattern maps are typically used throughout the development of all coefficients as a quality control device. The geographic boundaries shown were generated using the modified NCAR vector geography database described in Section 8.1.

2.1.4.1 Optimum Number of Coefficients for Predicting Parameters

The number of coefficients used to generate predicted parameters can influence the accuracy of the results. The higher the number of coefficients the more accurate the results. However, too many coefficients can result in computer storage and computational inefficiencies. Figures 2.3, 2.4, and 2.5 illustrate on a global scale the effect of predicting mean sky cover using 121, 169, and 225 coefficients.

The patterns in Figures 2.3, 2.4, and 2.5 are predicted mean cloud amounts between 80 degrees south and 80 degrees north for the winter season at 00 GMT. Coded patterns in the

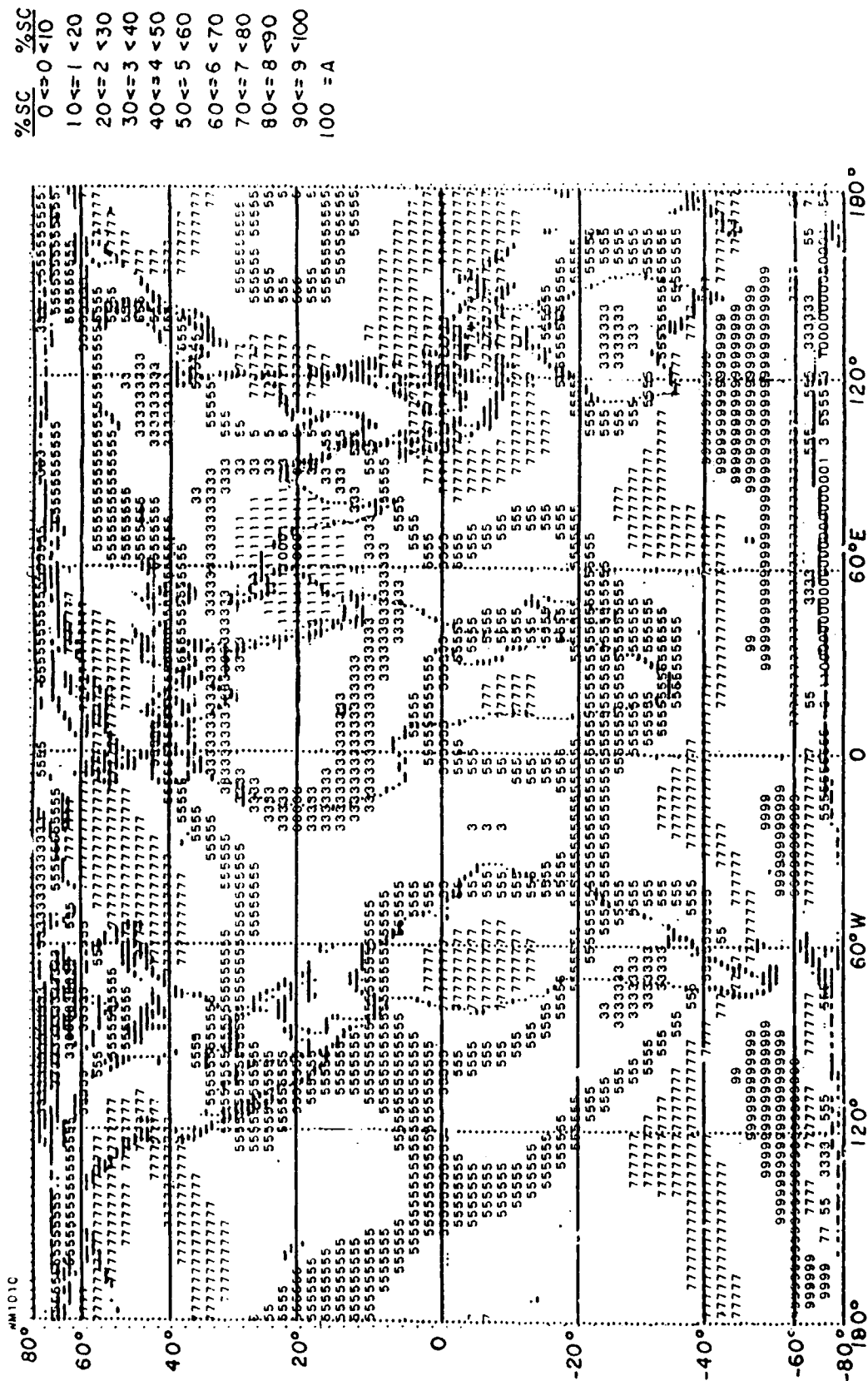


Figure 2.3. Pattern Map of Predicted Mean Sky Cover Using 121 Coefficients (10 x 10 Array Size). Original Data Were DOE Mean Cloud Cover for the Winter Season at 00 GMT.

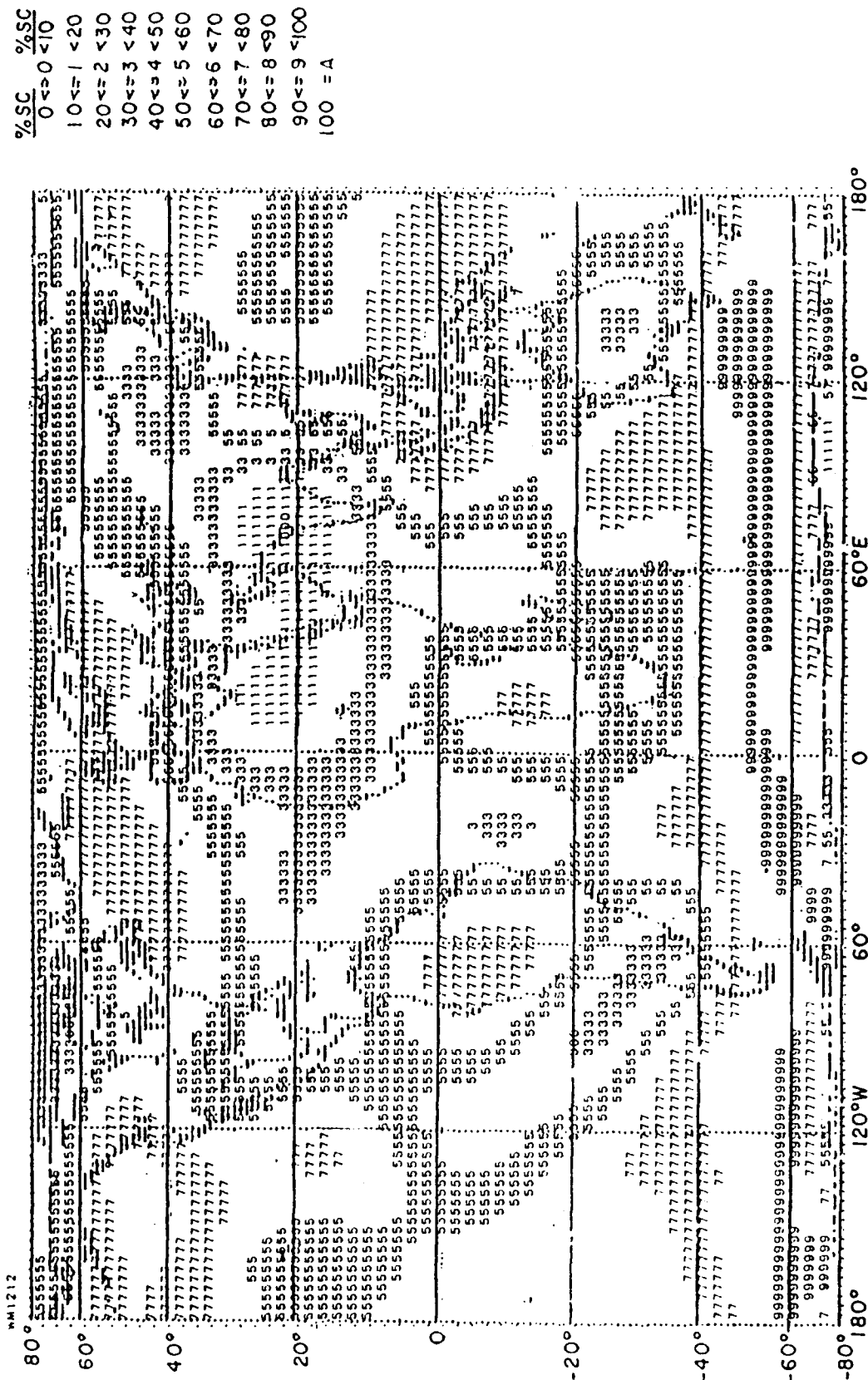


Figure 2.4. Pattern Map of Predicted Mean Sky Cover Using 169 Coefficients (12 x 12 Array Size). Original Data Were DOE Mean Cloud Cover for the Winter Season at 00 GMT.

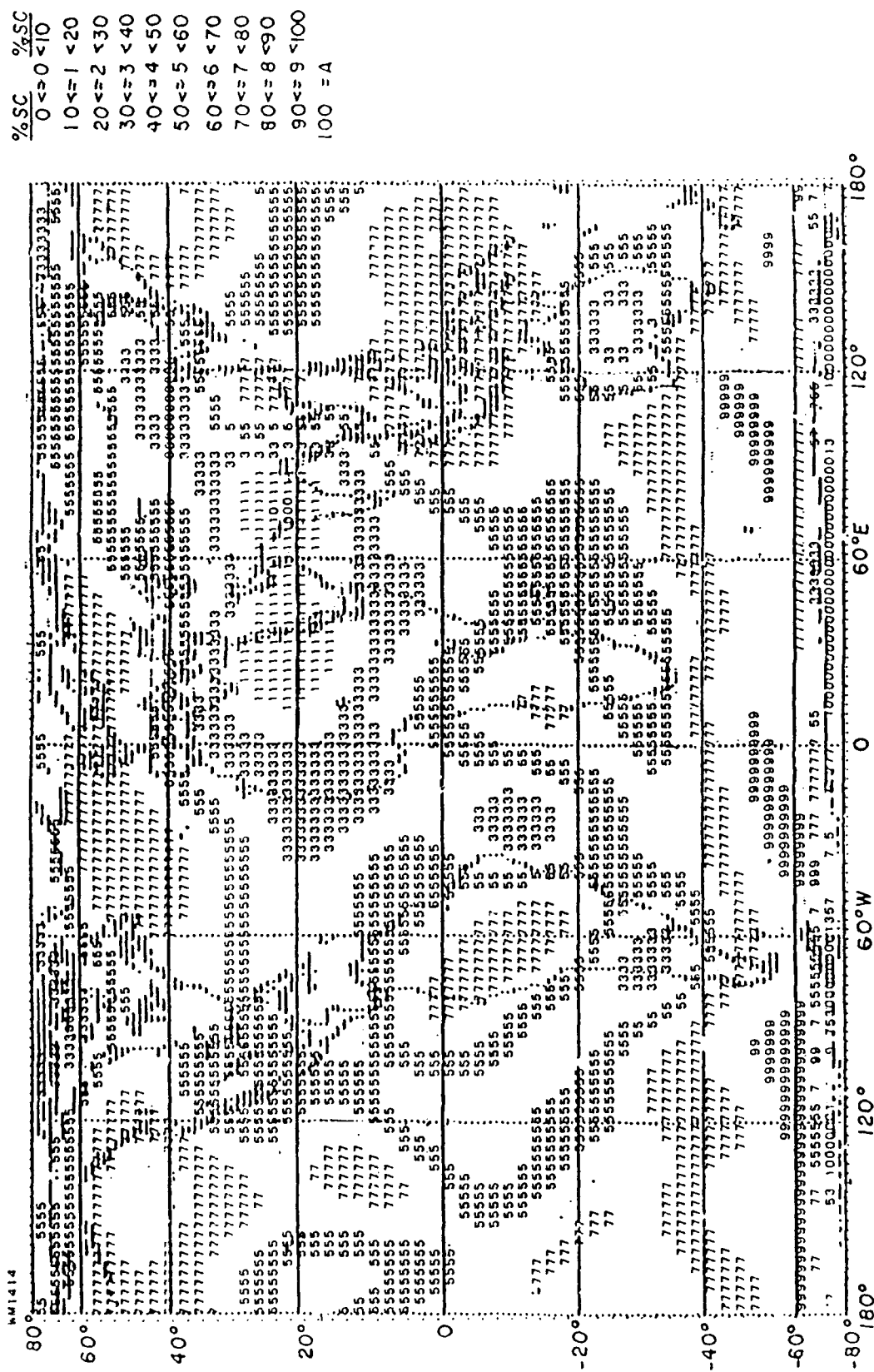


Figure 2.5. Pattern Map of Predicted Mean Sky Cover Using 225 Coefficients (14 x 14 Array Size). Original data were DOE Mean Cloud Cover for the Winter Season at 00 GMT.

maps such as 0,1,2, etc., represent the percent sky cover (SC) ($\%SC = \text{probability of mean sky cover} \times 100$) boundaries that are quantified in the index to the right of the maps. Even number codes are purposely left blank in order to enhance the patterns of overall sky cover amounts.

Figure 2.3 shows patterns of predicted sky cover amounts using 121 coefficients (10 x 10 array size). In Figure 2.4, a slight change over central Australia can be detected when going to a slightly larger number of coefficients (169 or 12 x 12 array size). Also, the patterns of predicted sky cover over the Sahara desert show a reduction of about 10 percent, while predicted sky cover amounts over the eastern Pacific Basin increased by a tenth of a percent.

Figure 2.5 patterns of predicted sky cover amount are generated using 225 coefficients (14 x 14 array size). These patterns are considered to be very close to actual DOE sky cover amounts. For example, patterns of mean sky cover amounts over central Australia have decreased from code 3 to 2 and those over the southwestern United States also show decreases from code category 5 to 4. These pattern maps can be compared with actual DOE mean cloud amounts shown in Figure 2.8 which in turn will be discussed in the following section.

Patterns of predicted mean correlation using 225 coefficients for the winter season at 00 GMT are illustrated

in Figure 2.6. In general, pattern codes for mean correlation of 5 or greater signify areas having sky cover distributions that are L, U, or J shaped. Areas having codes less than 5 have sky cover distributions that are generally flat or bell shaped.

Patterns of population are shown in Figure 2.7. In the development of coefficients for predicting populations, the values are converted to their natural logarithm. Therefore, the single code value representing a predicted population value in the pattern maps are gross estimates of population values. Nevertheless, one can observe that population values over land areas north of about sixty degrees south have reasonably robust values, but those over the ocean areas of the southern hemisphere are quite sparse.

2.1.4.2 Error Analysis of Predicted Mean Sky Cover

Figures 2.8 through 2.8.7 illustrate maps of DOE mean cloud cover for winter and summer together with maps of predicted mean cloud amounts, differences of mean cloud amounts minus predicted amounts, and population counts. The uncontroled data values are presented on linear cylindrical projections of the globe for detailed visual inspection of predicted values against given mean cloud amounts.

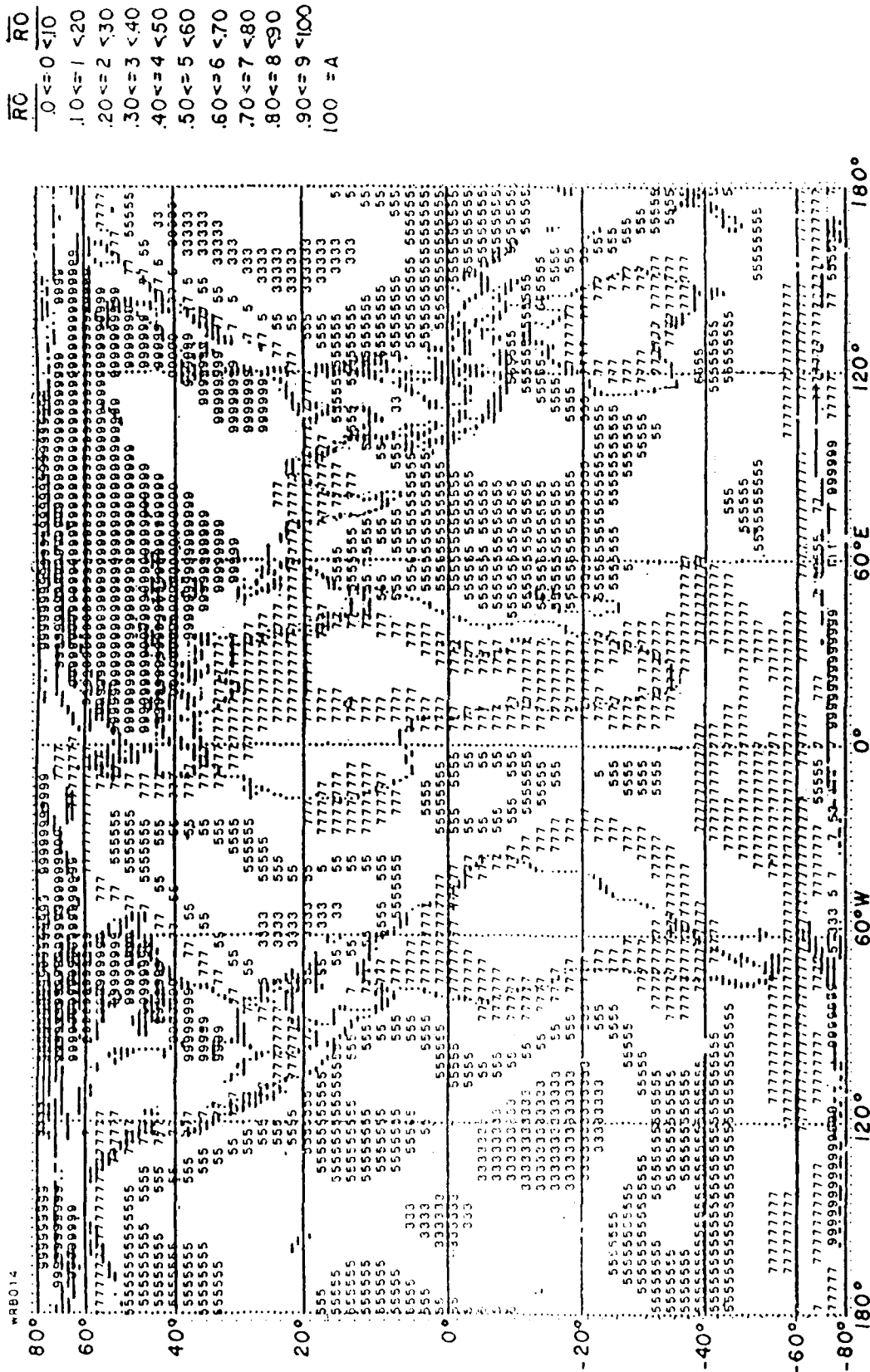


Figure 2.6. Pattern Map of Predicted Mean Correlation Using 225 Coefficients (14 x 14 Array Size). Original data were DOE Mean Cloud Cover and Variance for Winter at 00 GMT.

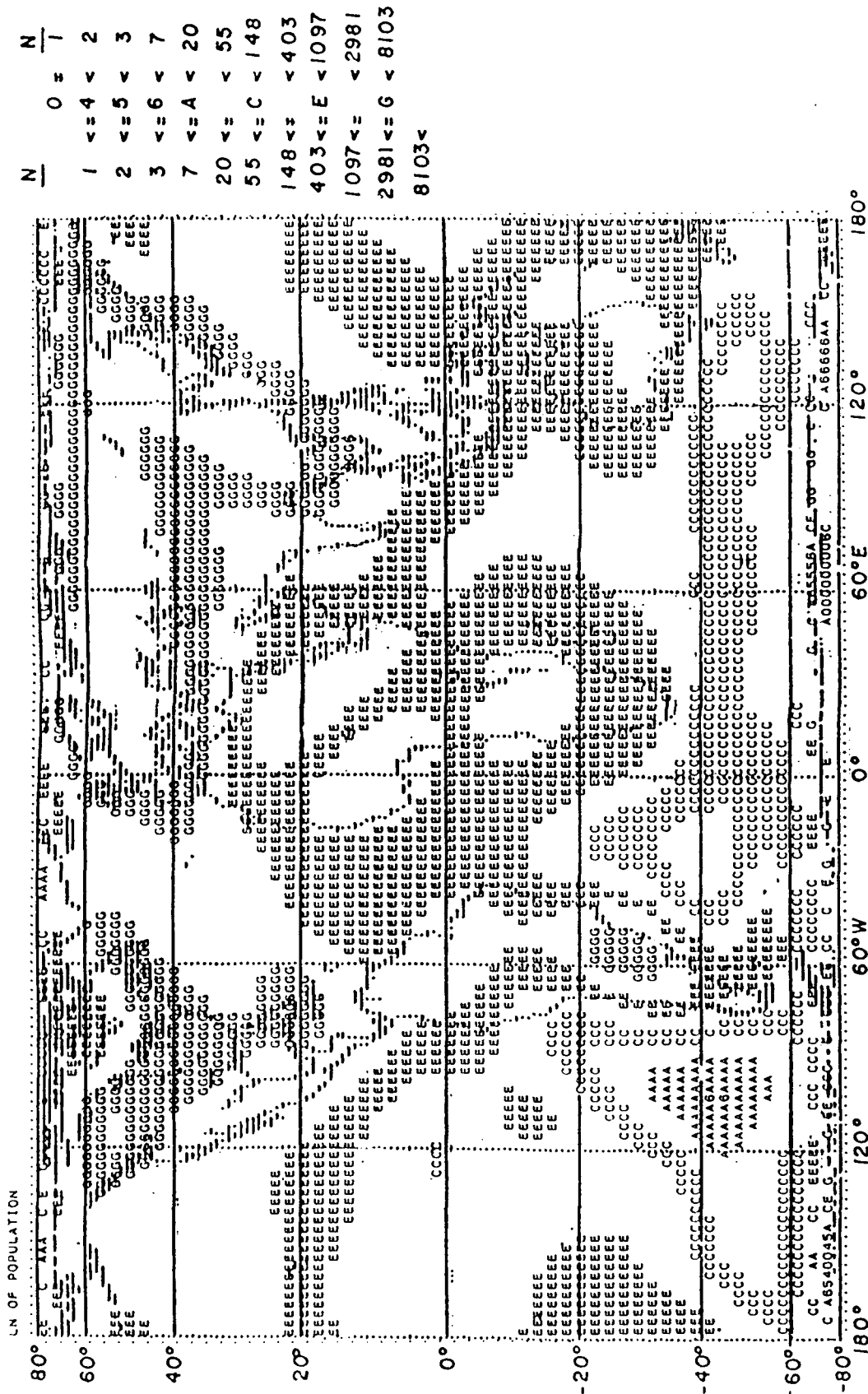


Figure 2.7. Pattern Map of Predicted Gross Population Using 225 Coefficients (14 x 14 Array Size). Original data were DOE Population Values of Mean Cloud Amounts for the Winter Season at 00 GMT.

Figure 2.8 represents values of DOE mean cloud cover for the winter season at 00 GMT. Blank areas represent missing data. Figure 2.8.1 shows predicted values of mean cloud cover for the same time period computed by using 225 coefficients (14 x 14 array size). (This array size will be the size chosen to develop coefficients for predicting cloud parameters.) Figure 2.8.2 is a display of the difference between the observed mean sky cover and the predicted mean sky cover. Figure 2.8.3 is included to show the approximate number of observations used in determining the mean cloud amounts in each DOE grid box. The actual population values were coded into one- and two-digit codes primarily for printing purposes. For example, if the population of an area was 978, the code number representing this value would be 978/100 or 9. Therefore, 9 represents populations between 900 and 999. Population values equal to or above 1,000 are coded 10. Code 0 represents values between 1 and 99. Figures 2.8.4 through 2.8.7 show the same sequence of events except that they are for the summer season at 00 GMT. As expected, rather large differences (plus or minus ten percent) in Figures 2.8.2 and 2.8.6 occur in areas bounded by high and low population counts. The weighting function in program W9090 causes values in areas of low population to be influenced by those areas with higher population. Population values for the summer case in Figure 2.8.7 are sparse in the southern hemisphere causing large differences in predicted sky cover amounts (plus or minus 30 percent). Many grid boxes over large areas in the

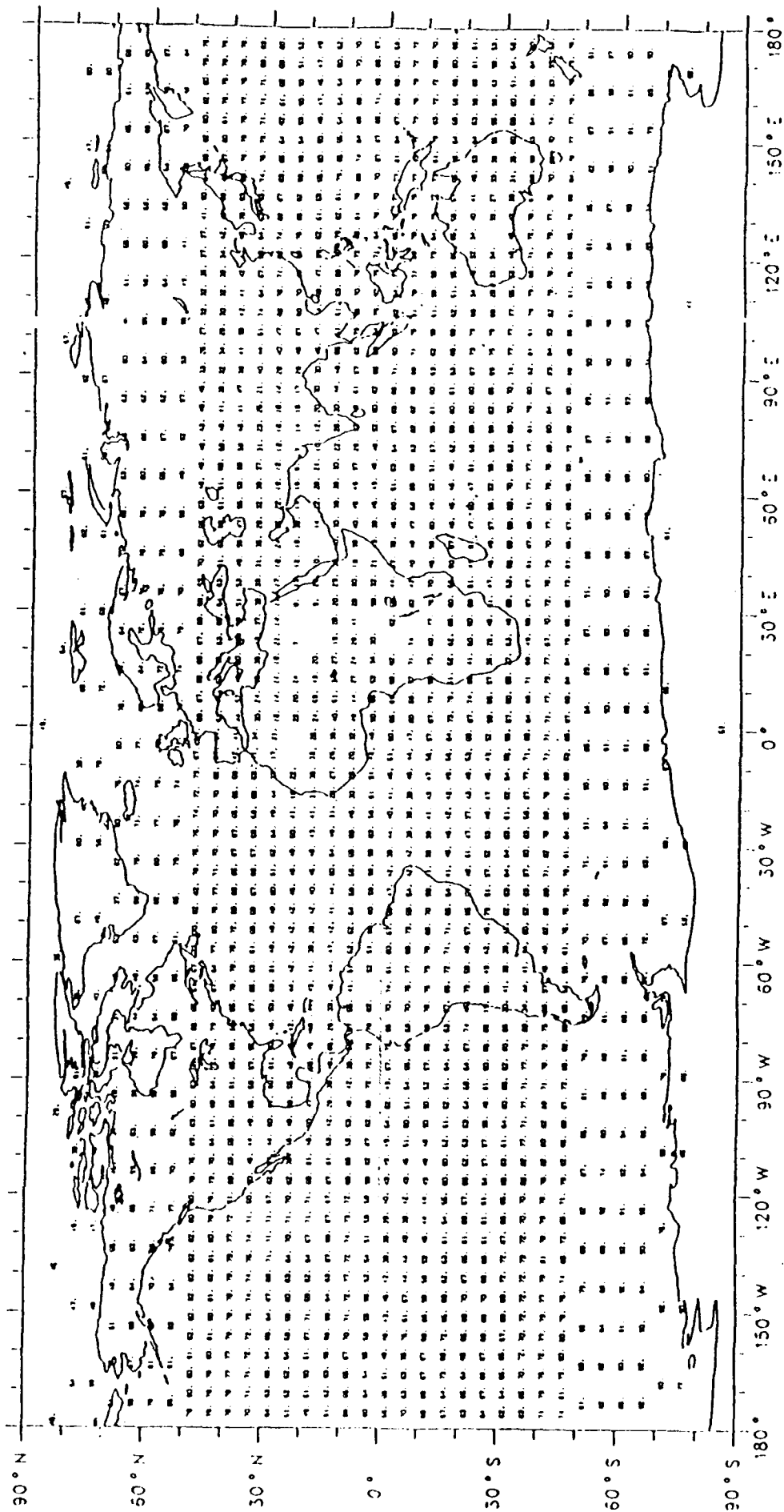


Figure 2.6. DOE Mean Cloud Cover for the Winter Season at 00 GMT.

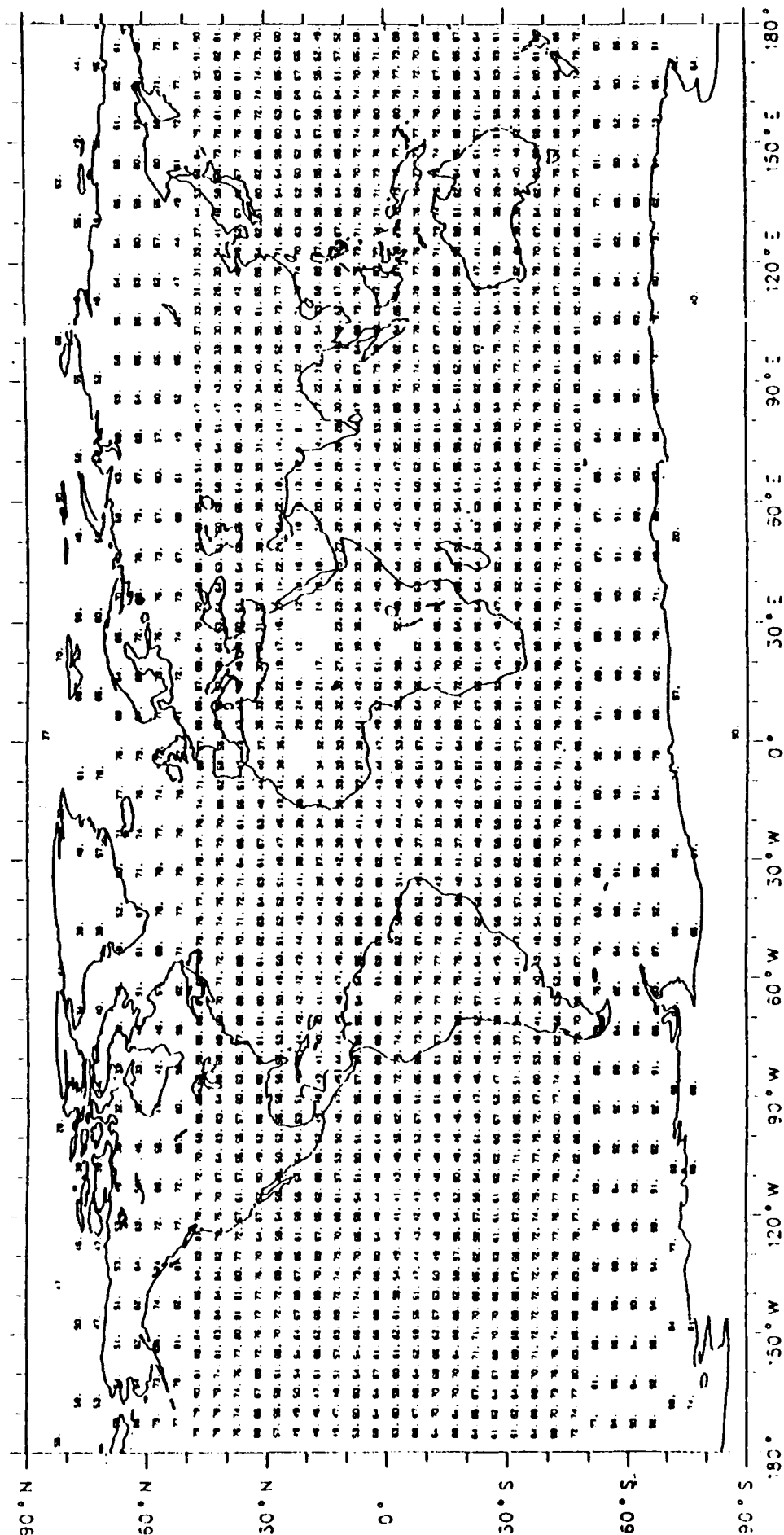


Figure 2.8.1. Predicted Mean Cloud Cover Using 225 Coefficients Derived from DOE Mean Cloud Cover for the Winter Season at 00 GMT.

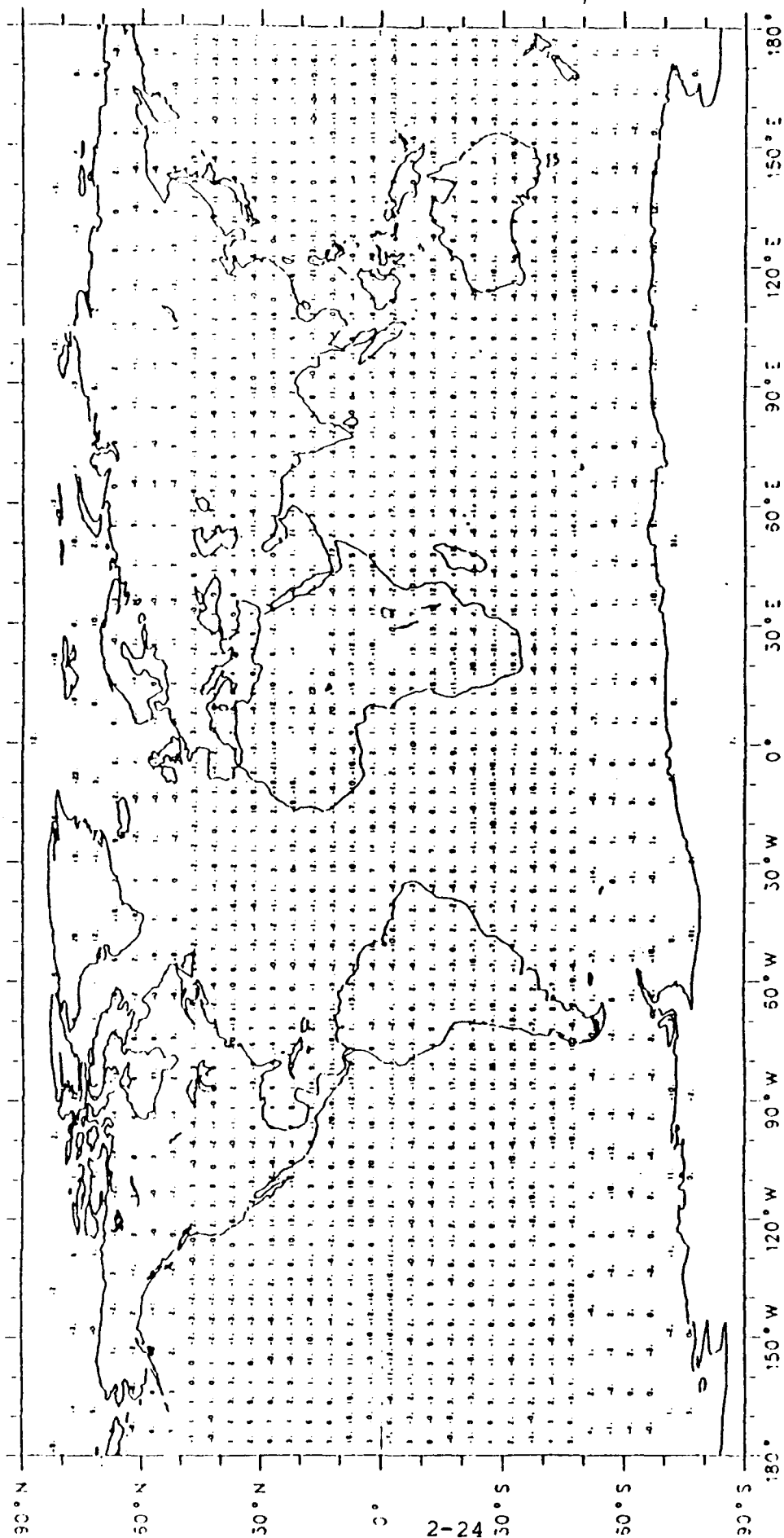


Figure 2.8.2. Difference Map of DOE Mean Cloud Cover for the Winter Season at 00 GMT
Minus Predicted Cloud Cover.

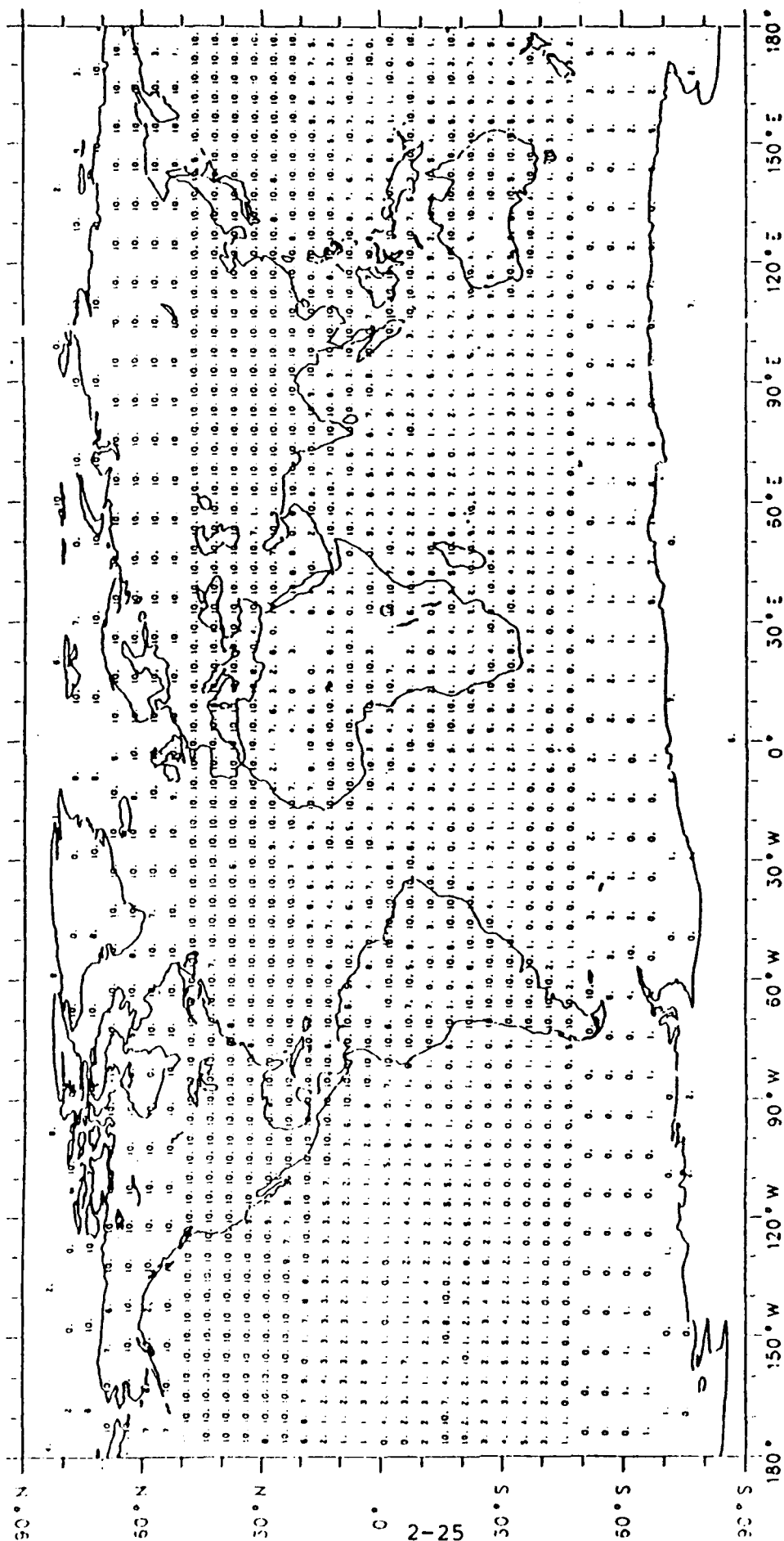


Figure 2.8.3. Coded DOE Population Map of Mean Cloud Cover for the Winter Season at 00 GMT. Gross Estimates of Actual Population Values are Computed as Code * 100. Hence, Code 9 is a Population Greater Than or Equal to 900 but Less Than 1,000. Blank Areas Reflect Missing Data. Code 10 Represents Population Values Greater Than or Equal to 1,000.

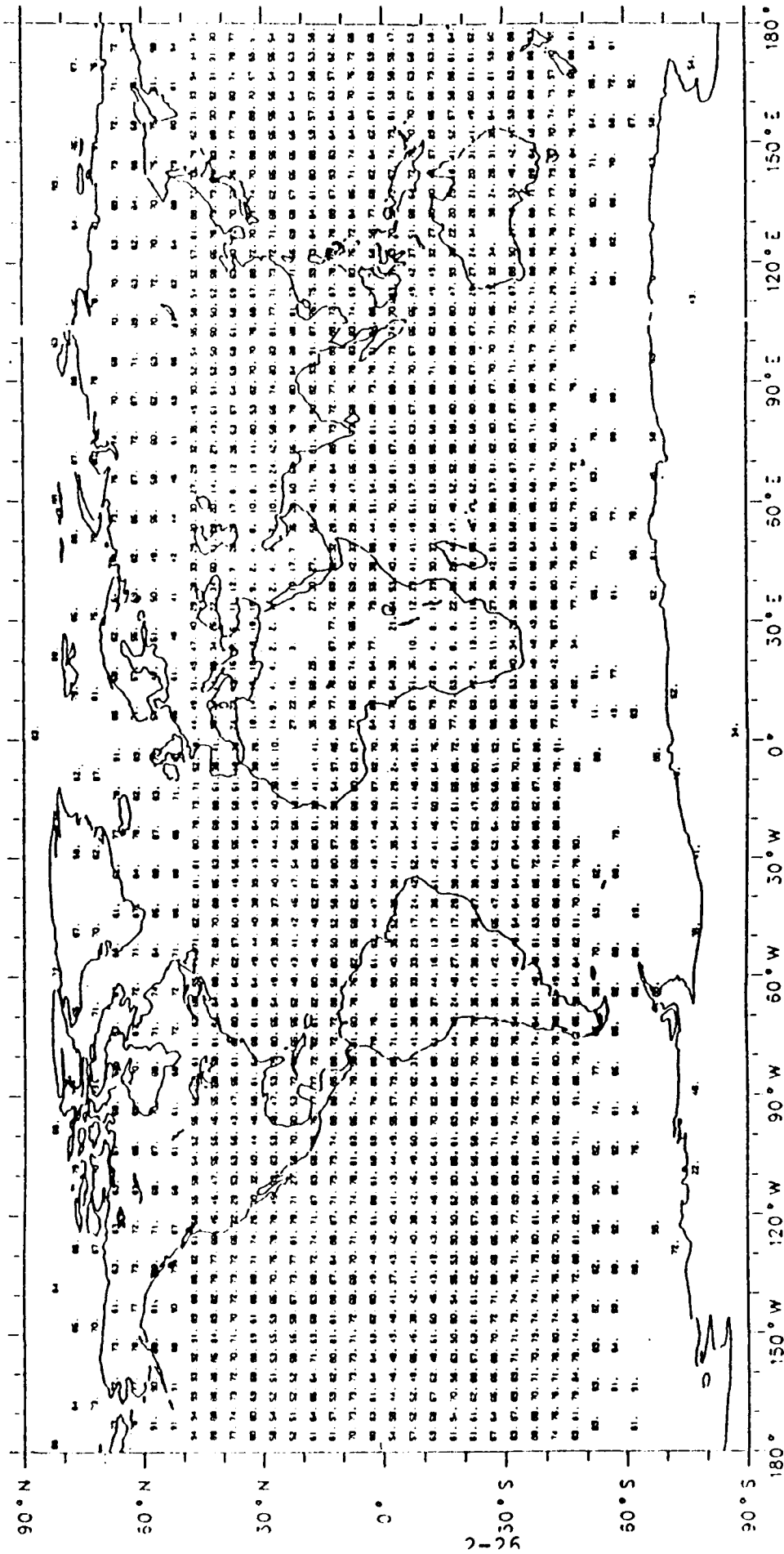


Figure 2.8.4. DOE Mean Cloud Cover for the Summer Season at 00 GMT.

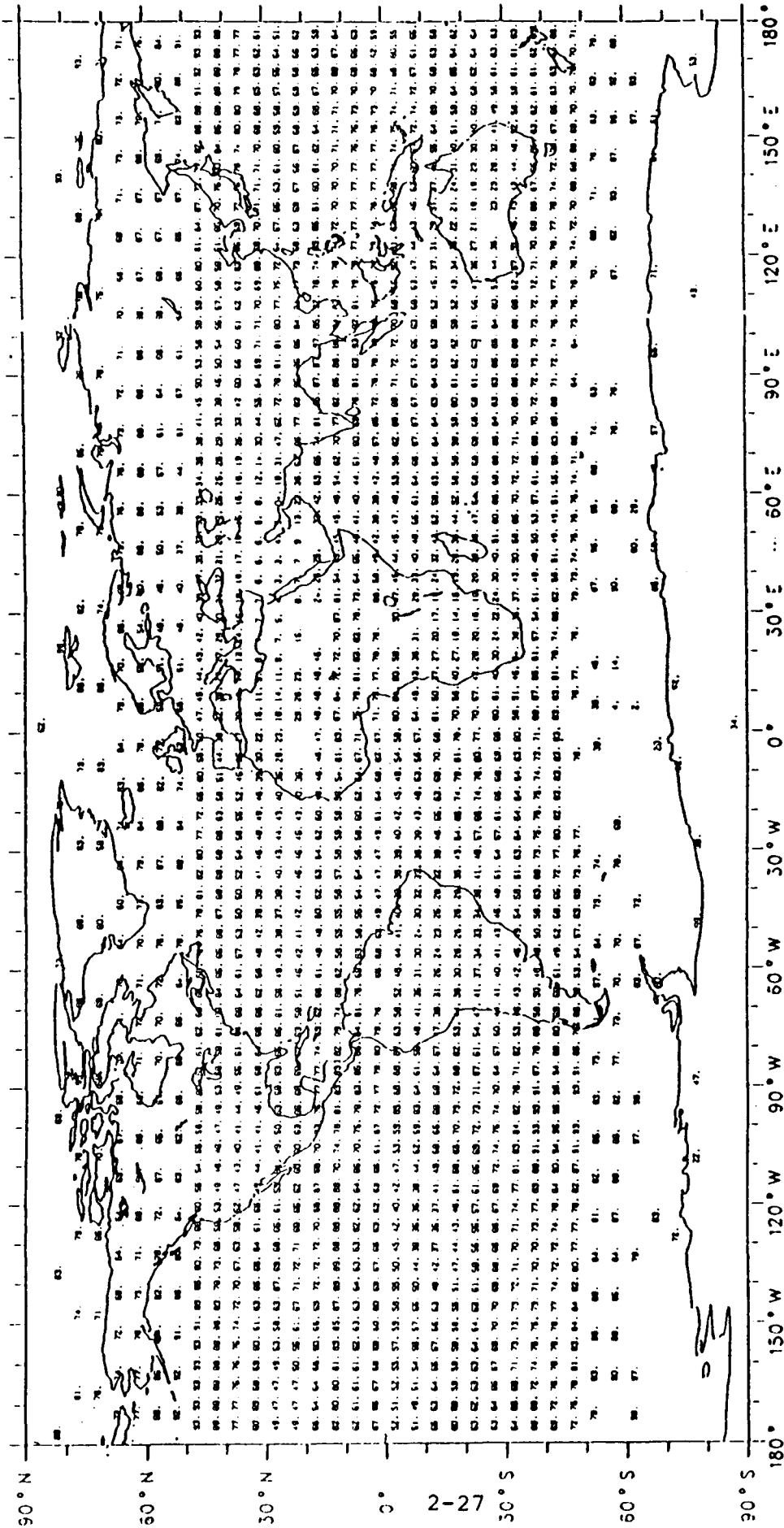


Figure 2.8.5. Predicted Mean Cloud Cover Using 225 Coefficients Derived from DOE Mean Cloud Cover for the Summer Season at 00 GMT.

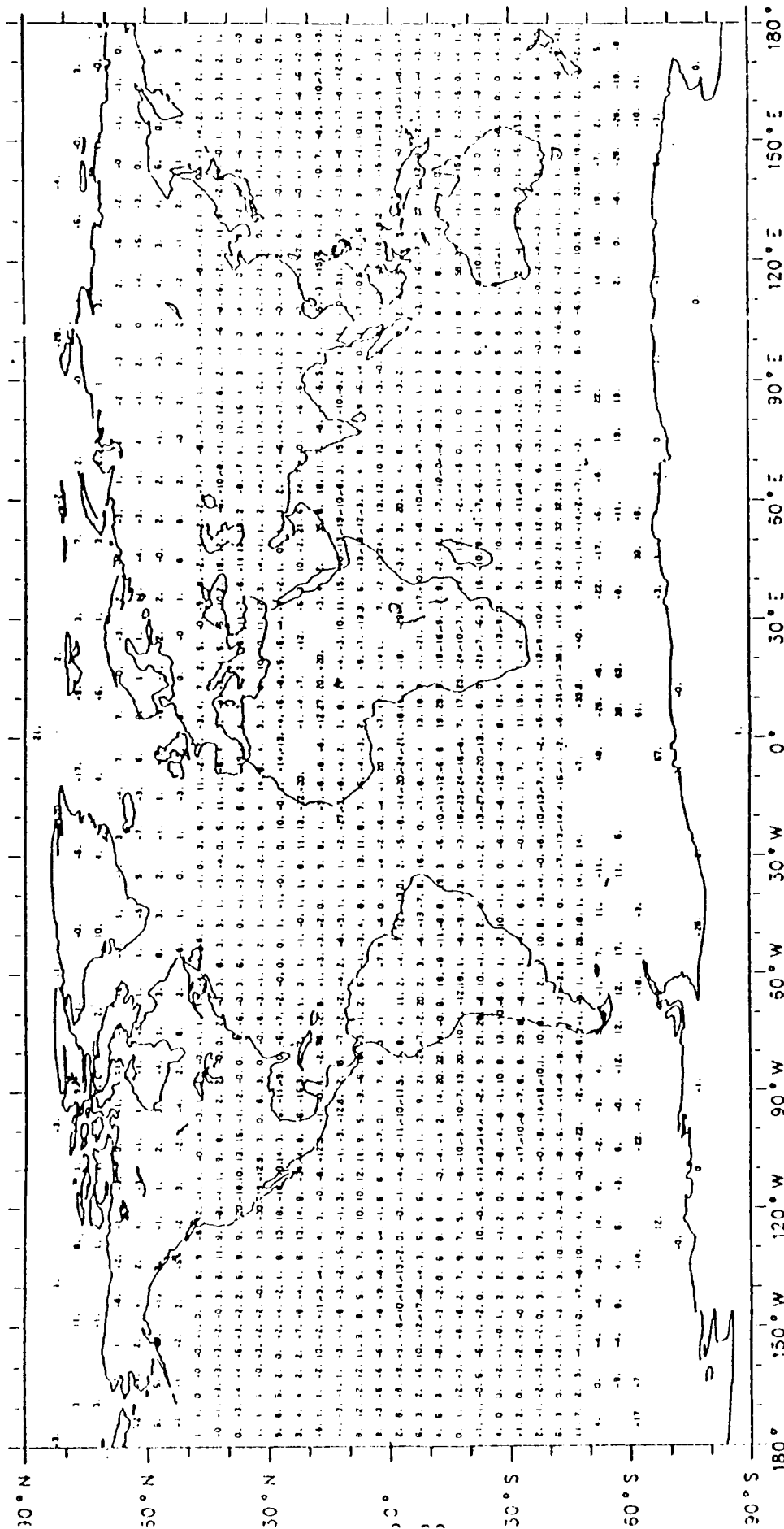


Figure 2.8.6. Difference Map of DOE Mean Cloud Cover for the Summer Season at 00 GMT
Minus Predicted Mean Cloud Cover.

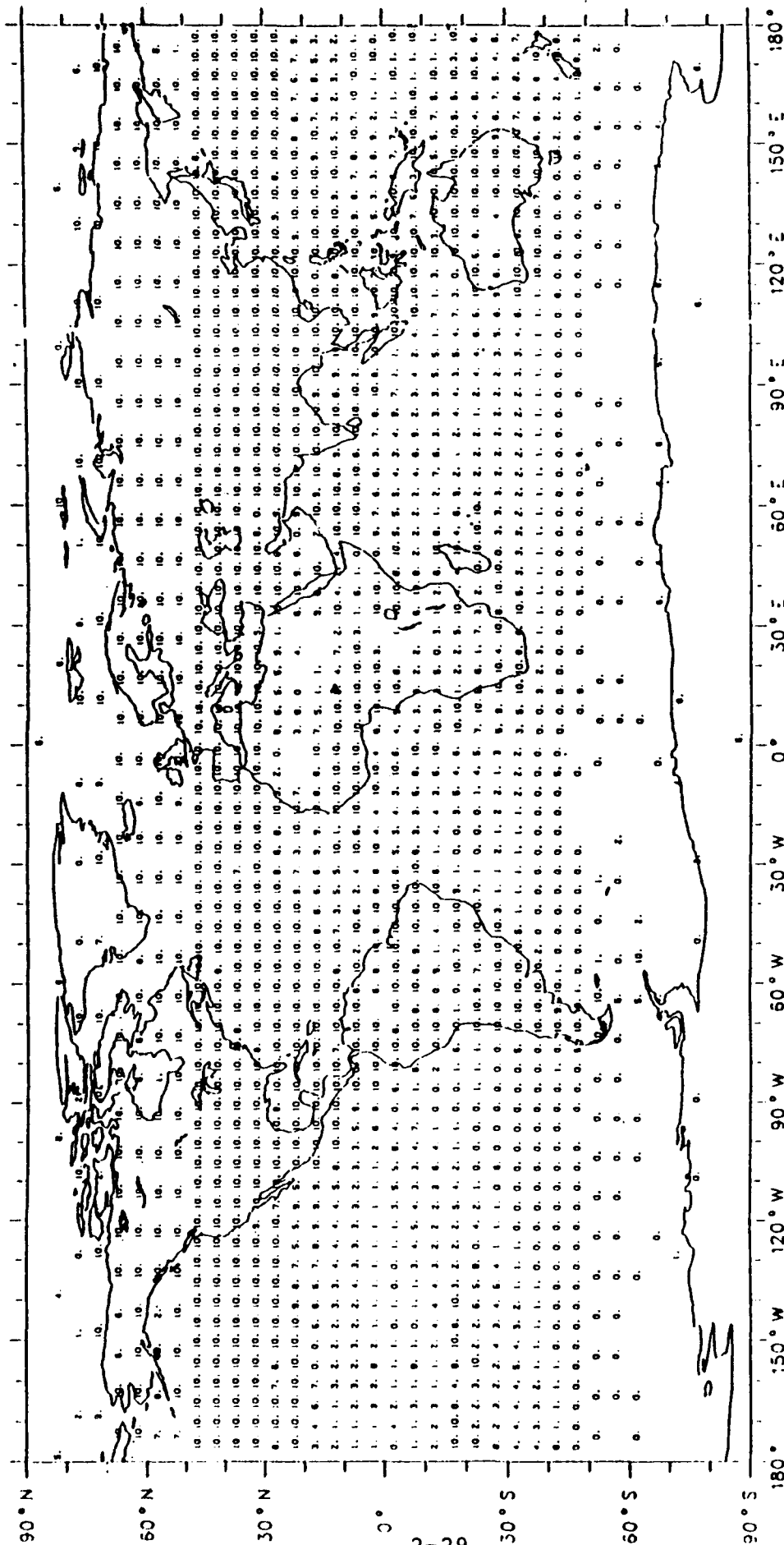


Figure 2.8.7. Coded DOE Population Map of Mean Cloud Cover for the Summer Season at 00 GMT. Gross Estimates of Actual Population Values are Computed as Code * 100. Hence, Code 9 is a Population Greater Than or Equal to 900 but Less Than 1,000. Blank Areas Reflect Missing Data. Code 10 Represents Population Values Greater Than or Equal to 1,000.

southern hemisphere have no data at all. This is also true of smaller less populated areas over the Amazon, Sahara Desert, Central Africa, and Central Australia.

2.2 Processing Interannual Variance of Mean Sky Cover

Sky cover amount over a given area of the globe for a particular month can vary somewhat from year to year. Therefore, in a system of automated global cloud climatology the magnitude of the interannual variation of sky cover amount is of significant interest for defining error conditions of predicted sky cover. The processing of the DOE cloud database to provide the interannual variation of sky cover from year to year is discussed next.

2.2.1 Program FETC2

As shown in Figure 2.8.8, a return to the DOE cloud data tape was necessary to extract monthly mean cloud cover data to compute interannual variations. Program FETC2, invoked by XEQFE2, extracts from file number 2 the January monthly mean cloud cover over land areas that are found under map group numbers 102 through 233. A second run was then made to extract the same parameter for the grid points over the ocean areas. These were found in file 6 under map group numbers 1340 through 1699. The extracted data are labeled on disk

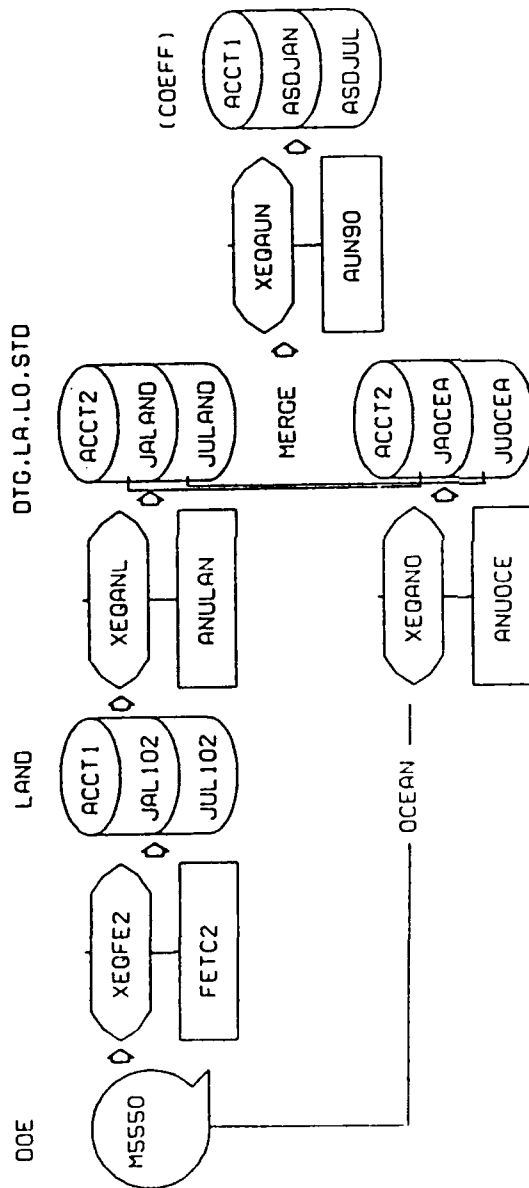


Figure 2.8.8. Data Flow Configuration for Processing Interannual Variance of Sky Cover Using DOE Data.

file as JAN102 for January land data and JUL102 for July land data.

2.2.2 Program ANULAN, AUNOCE

Program ANULAN reads the monthly mean values extracted from FETC2 and sums them for each grid box. The sum of the squares of the percentages are also computed and at the completion of the process the standard deviation is computed for each grid box using the basic formula below.

$$v = \frac{\sum xm^2}{N} - \left(\frac{\sum xm}{N} \right)^2 \quad (2.7)$$

$$SD = \pm \sqrt{V} \quad (2.8)$$

where

XM is the monthly mean cloud amount

N is the number of observations

V is the variance

SD is the standard deviation

A date time group DTG, latitude-longitude reference, and positive standard deviation are stored as files JALAND and JULAND for January and July respectively. Program ANUOCE handles the ocean data in the same manner. January ocean data are stored as JAOCEA and July as JUOCEA. At the completion of this process, the land data are merged with the ocean and the JALAND, JULAND files are purged.

2.2.3 Program AUN90

Finally, program AUN90, a downgraded version of W9090 with the weighting scheme removed, uses only the interannual standard deviations read from files JAOCEA and JUOCEA to compute coefficients for predicting monthly interannual deviation of cloud amount. The coefficients for January positive standard deviations are stored under the name ASDJAN and JULY as ASDJUL.

Figure 2.8.9 shows the patterns of predicted interannual standard deviations of mean sky cover derived from coefficients using the 11 DOE Januarys.

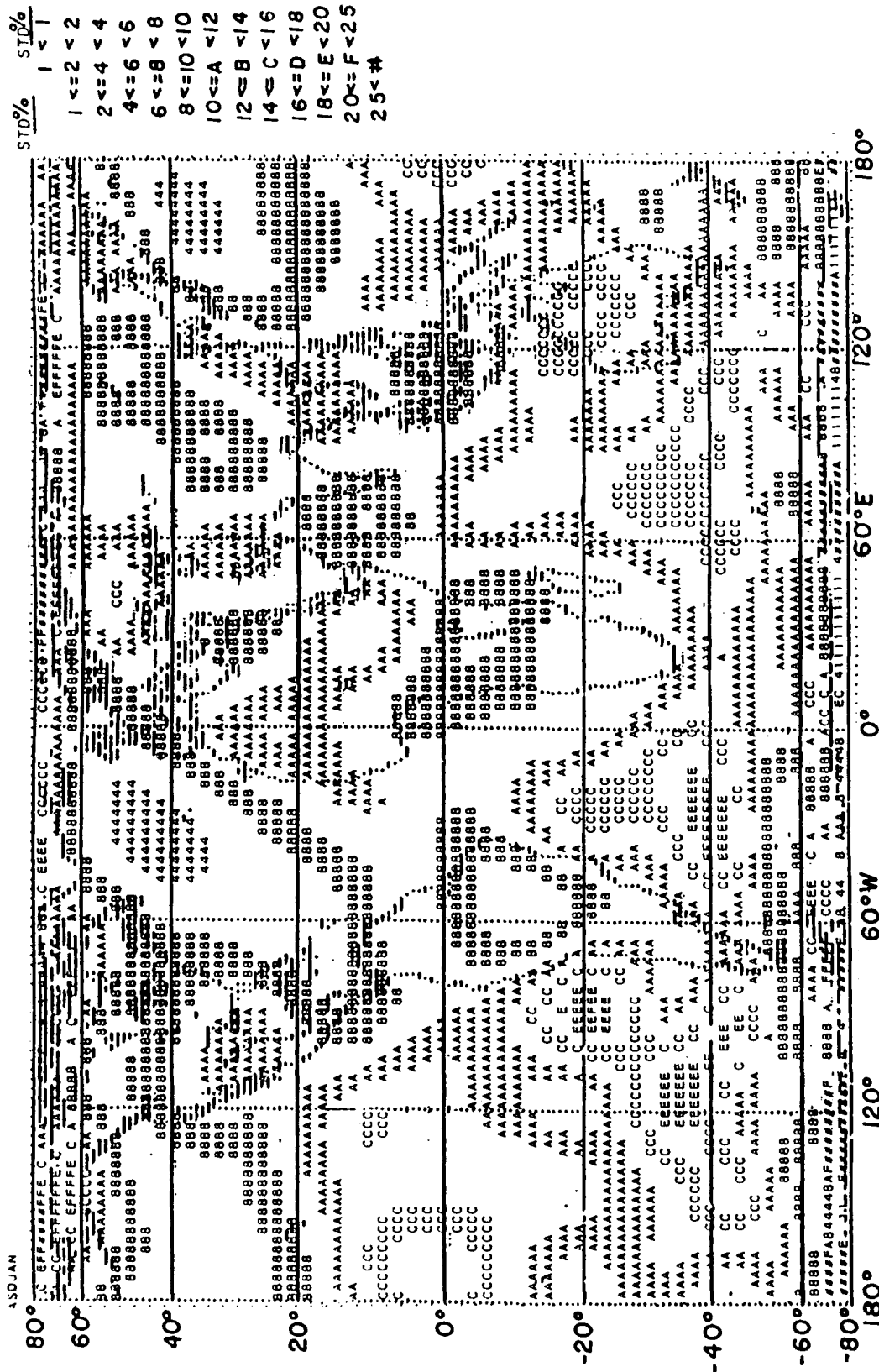


Figure 2.8.9. Pattern Map Showing Predicted Interannual Standard Derivation of Mean Cloud Cover from DOE Data.

3.0 NIMBUS 7 SATELLITE CMATRIX CLOUD DATA

According to Hwang, et al. (1988), the primary merits of the NIMBUS-7 CMatrix Cloud data set are 1) observed radiances used for this cloud data set are from the same instruments used over six continuous years between 1979-1985, so it represents the most homogeneous satellite derived cloud data set available; 2) the cloud data set is efficiently retrieved and stored, so that the analysis of global cloud distribution over climatological time scales is feasible; and, 3) daily as well as monthly averages and variances of cloud data are presented together with a correlation coefficient surface temperature archive. Because of these merits, this cloud database was acquired to evaluate its usefulness in the development of the automated global cloud climatology.

NIMBUS earth viewing satellites are placed in sun synchronous orbits. They provide global coverage in every 24-hour period at noon and midnight.

The NIMBUS 7 satellite CMatrix cloud data was received on seven separate magnetic tapes. The data on these tapes consist of more than 100 parameters that are archived twice a day at noon and midnight. The parameters archived are daily means of percent coverage of total, low, middle, and high cloud layer amounts, clear air radiances, and their spatial and temporal variances. These parameters, identified by

number, are computed for 2070 NIMBUS 7 earth radiation budget (ERB) target areas. Each target represents an area of approximately 500 x 500 kms. Typical spatial distribution of data within the 2070 ERB target areas can be seen in Figure 3.0. In this figure every other south-to-north satellite pass is purposely left blank to show the actual data positioning within each satellite swath. The numbers in the figure represent data observed in Greenwich Mean Time (GMT) during ascending nodes (noon time). The data set has a six-year period of record starting in April 1979 through March 1985.

3.1 Processing of NIMBUS 7 CMatrix Cloud Data

Figure 3.1 illustrates the flow of events used to process the NIMBUS 7 CMatrix cloud databases.

3.1.1 Program EXTRAK

Procedure XEQFET calls on program EXTRAK to extract from a given CMatrix data tape the daily maps of parameters selected by the parameter number defined in the CMatrix tape user documentation by Wellemeyer (1986). A list of the tape numbers assigned to the acquired CMatrix data tapes and their archived period of record are listed in Table 3.0 below.

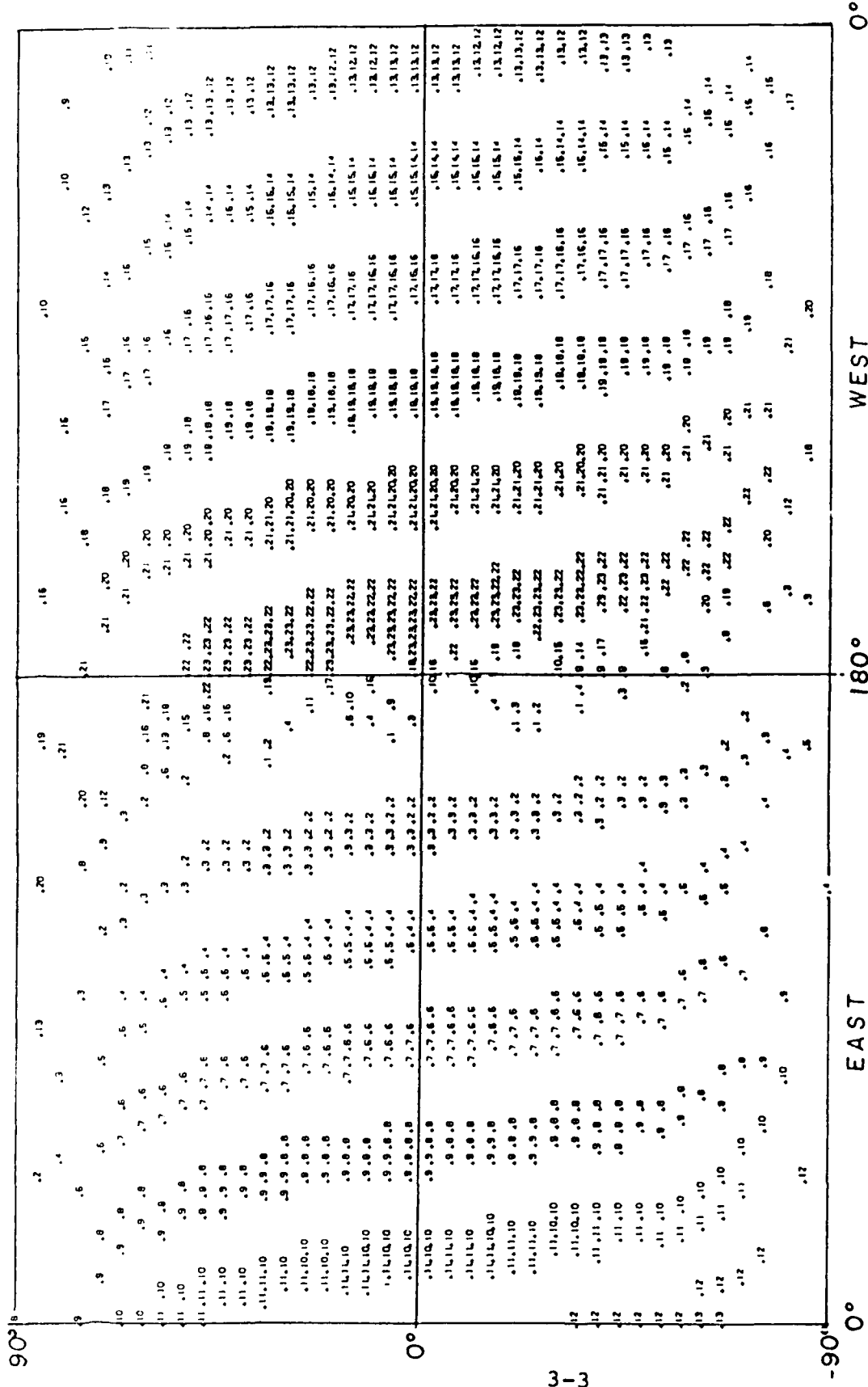


Figure 3.0. A Track Map of NIMBUS 7 CMatrix Data Dissemination Over 2,070 ERB Target Areas. Every Other Orbital Swath is Purposely Left Blank to Show the South-to-North Noon Time Satellite Coverage. Numbers in the Figure Represent GMTs of Data Observations.

TABLE 3.0**CMatrix Tape Numbers and Archived Period of Record**

Tape Number	Period of Record
M8197	Apr 79 - Oct 79
M8196	Nov 79 - Oct 80
M8199	Nov 80 - Oct 81
M8202	Nov 81 - Oct 82
M8204	Nov 82 - Oct 83
M8407	Nov 83 - Oct 84
M8349	Nov 84 - Mar 85

A list of the selected parameters extracted from tape are listed by number code in Table 3.1.

TABLE 3.1**Parameter Definition Given Selected CMatrix
Parameter Number Code**

Number Code	Definition
1	Recommended total percent cloudiness ascending
2	RMS of recommended total percent cloudiness ascending
23	Mean subtarget area population ascending
25	GMT (minutes) ascending
55	Total percent cloudiness descending
56	RMS of total percent cloudiness descending
75	Mean suntarget area population descending
77	GMT (minutes) descending

The extracted daily parameter maps for all months from a single CMatrix tape are stored on a second set of tapes for further processing.

3.1.2 Program SATSUM

Program SATSUM is utilized to sum the daily data stored on the new tapes for a selected month and parameter code. The program uses a navigation table called TABDA, published by Wellemeyer (1986), to navigate the 2070 NIMBUS 7 ERB target areas. Daily sums of 4 parameters over individual months for six Januarys and Julys are computed and stored in binary form on files named JANSAl and JANSA2, etc. The four parameters summed are 1) mean cloud cover, 2) mean cloud cover squared, 3) GMTs, and 4) number of observations.

3.1.3 Program NORSAT

Program NORSAT reads in the binary data summation files from SATSUM and computes the mean and variance of cloud cover in each ERB target area for both ascending and descending nodes using 3.01 and 3.02.

$$\bar{M} = \frac{\sum M}{N} \quad (3.01)$$

$$V = \left[\frac{\sum M^2}{N} - (\bar{M})^2 \right] \quad (3.02)$$

where

ΣM = Sum of daily mean cloud cover, ascending
or descending node, for a given month
within an ERB target area

\bar{M} = Mean cloud cover

N = Population count

ΣM^2 = Sum of the squares of the daily mean
cloud covers

V = Variance of mean cloud cover

A date-time group, latitude-longitude locator, mean cloud cover, standard deviation, mean correlation, and population for each ERB target area are then written to permanent files in a format like that described in Section 2.1.2. Names assigned to the files are SATJA1, SATJA2, SATJU1, and SATJU2 which represent satellite January ascending (1) and descending (2) and July ascending (1) and descending (2), respectively.

3.1.4 Program W9090

At this point, program W9090 is executed to derive the sets of 225 (14 x 14 array size) coefficients for predicting mean sky cover, mean correlation, standard deviation, and population. The names of these files like SATMWA are the coefficients from satellite (SAT) for mean cloud cover (M) in the winter (W) and noon ascending (A) node. The coefficients

stored for descending (D) node data would be labeled SATMWD, SATSWD, SATRWD and SATPWD.

3.1.5 Program MAPDIS

Program MAPDIS inputs the coefficients generated by program W9090 to produce global pattern maps for data quality control purposes. An example of the pattern map produced showing predicted mean cloud cover from satellite data for January at local noon is portrayed in Figure 3.2.

The corresponding pattern map of predicted mean correlation from satellite data is shown in Figure 3.3.

Tape M12096 is a permanent backup tape for programs, data, and coefficients. A floppy disk can also now be utilized to store the coefficients for use on smaller personal computers.

3.2 Cloud Cover Correlation Structure Using CMatrix Cloud Data

Program COVSPA, not shown in Figure 3.1, processes the CMatrix cloud amount data to investigate the cloud cover correlation structure as seen from a down-looking satellite. The CMatrix data processing procedure for deriving cloud cover

%SC %SC
0 <= 0 <10
10 <= 1 <20
20 <= 2 <30
30 <= 3 <40
40 <= 4 <50
50 <= 5 <60
60 <= 6 <70
70 <= 7 <80
80 <= 8 <90
90 <= 9 <100
100 = A

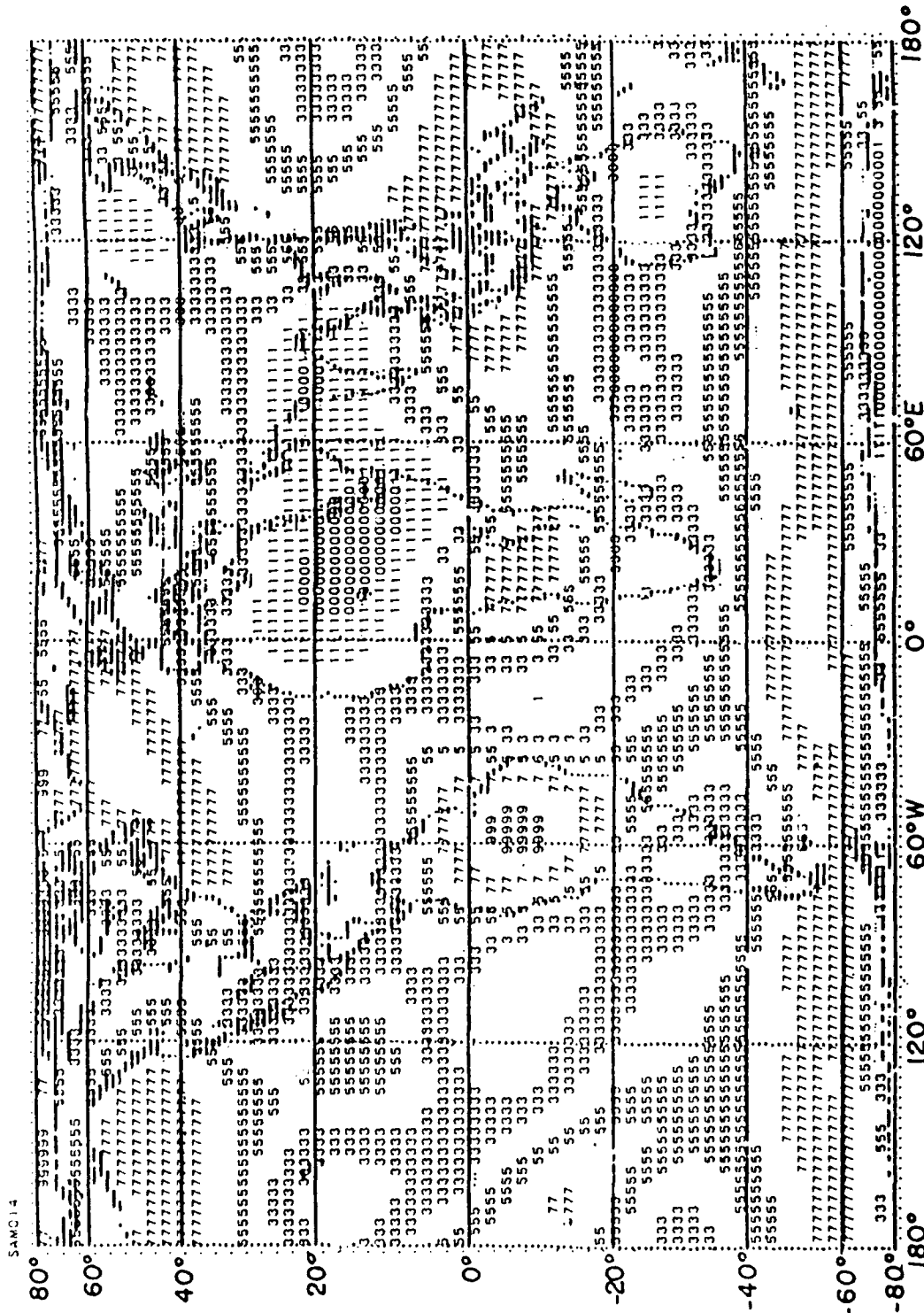


Figure 3.2. Pattern Map of Predicted Mean Sky Cover Using 225 Coefficients (14 x 14 Array Size). Coefficients Were Derived from NIMBUS 7 CMatrix Mean Cloud Cover for the Month of January at Local Noon.

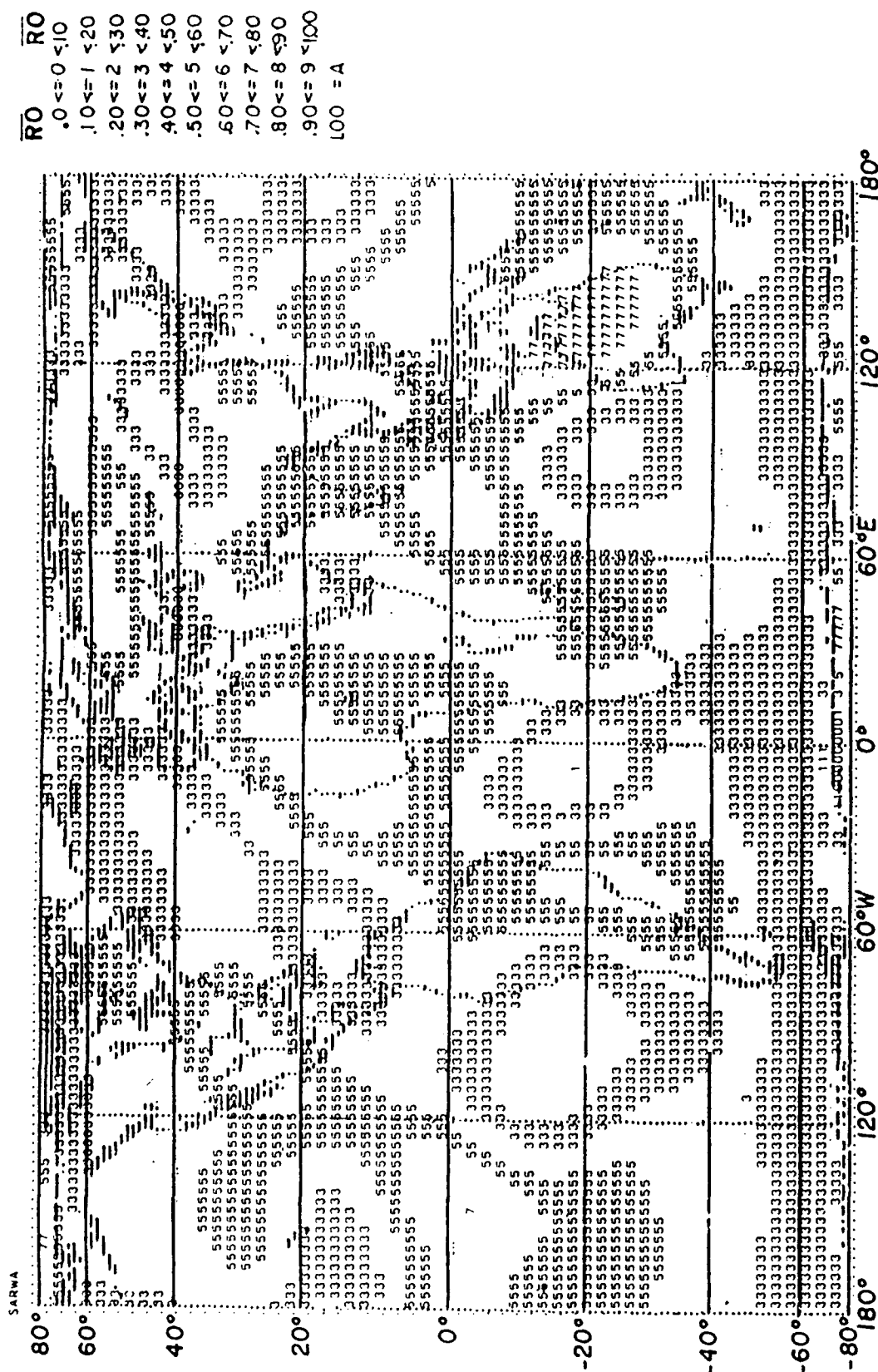


Figure 3.3. Pattern Map of Predicted Mean Correlation ($\overline{R_0}$) Using 225 Coefficients (14 x 14 Array Size). Coefficients Were Derived from NIMBUS 7 CMatrix Mean Cloud Cover and Variances for January at Noon.

correlations from a down-looking satellite is similar to that described in Section 7.2.3 which deals with sky cover correlation structure as seen from up-looking whole sky photographs.

The processing begins by transforming the CMatrix percent cloud amounts in each ERB target area into dichotomous values of partly cloudy, defined by less than or equal to 50 percent cloudiness, and mostly cloudy, defined by greater than 50 percent cloudiness. The given cloud condition over a central ERB target, centered within a 60-degree latitude x 90-degree longitude area, is then compared with the cloud condition in each surrounding ERB target. Results are summed and tabulated in contingency tables and at the completion of the tabulations the fourfold table values are used to compute a tetrachoric correlation for each target. Finally, the results are mapped and contoured for visual interpretation.

The program was utilized to generate cloud correlation structure maps over many areas of the globe for the month of April. Figure 3.4 shows one example of typical results. This case was chosen for presentation because of its proximity to Columbia, MO, where patterns of correlation structure of sky cover from up-looking whole sky photos are presented in Section 7.2.3.

As shown in Figure 3.4, the .8 correlation contour near the center position appears nearly isotropic. Moving further away from the center, however, the contours of lesser correlation appear more elliptical. Negative correlations are represented by broken contours. Zero correlation occurs at or about plus or minus 10.5 degrees or 630 nautical miles from the center.

These types of correlation structures derived from sensor data aboard satellite earth viewing systems may be helpful in defining global cloud motions in future cloud simulation systems.

3.3 Processing CMatrix Data for Interannual Variances of Mean Cloud Cover

Data flow used to process the NIMBUS 7 CMatrix data for monthly interannual variance of mean cloud cover is depicted in Figure 3.5.

3.3.1 Program NUSAT

Program NUSAT through procedure XEQNUS, utilizes the CMatrix monthly summation files described in Section 3.1.2 to compute the monthly interannual variance of mean cloud cover from satellite data. File TABDA is inputted first to navigate the 2070 CMatrix ERB target areas. The files of summed

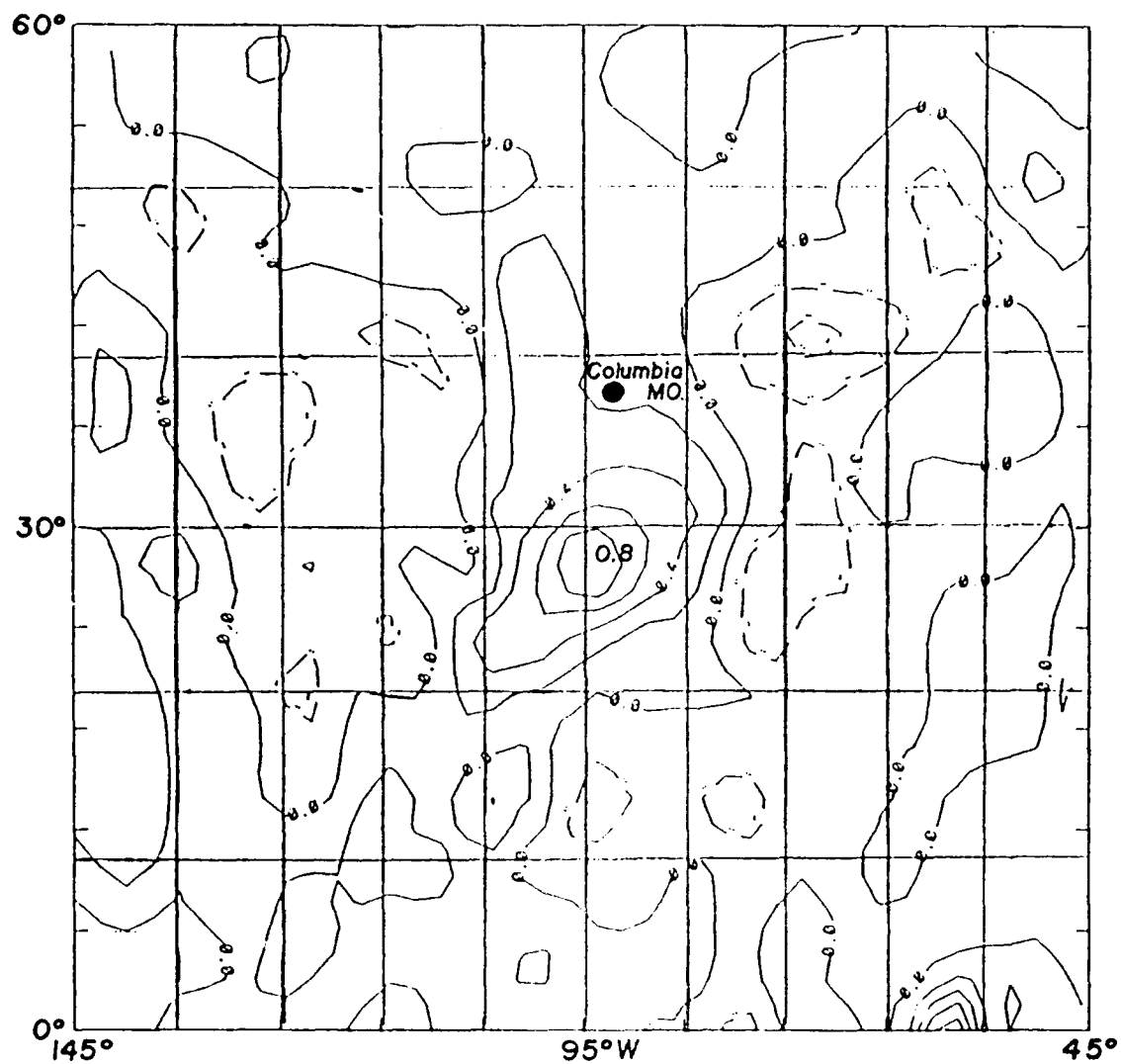


Figure 3.4. Contours of Cloud Cover Correlation Structure from NIMBUS 7 CMATRIX Total Percent Cloudiness Ascending Node (Noon) for Six Aprils. (Broken Contours Represent Negative Correlations.)

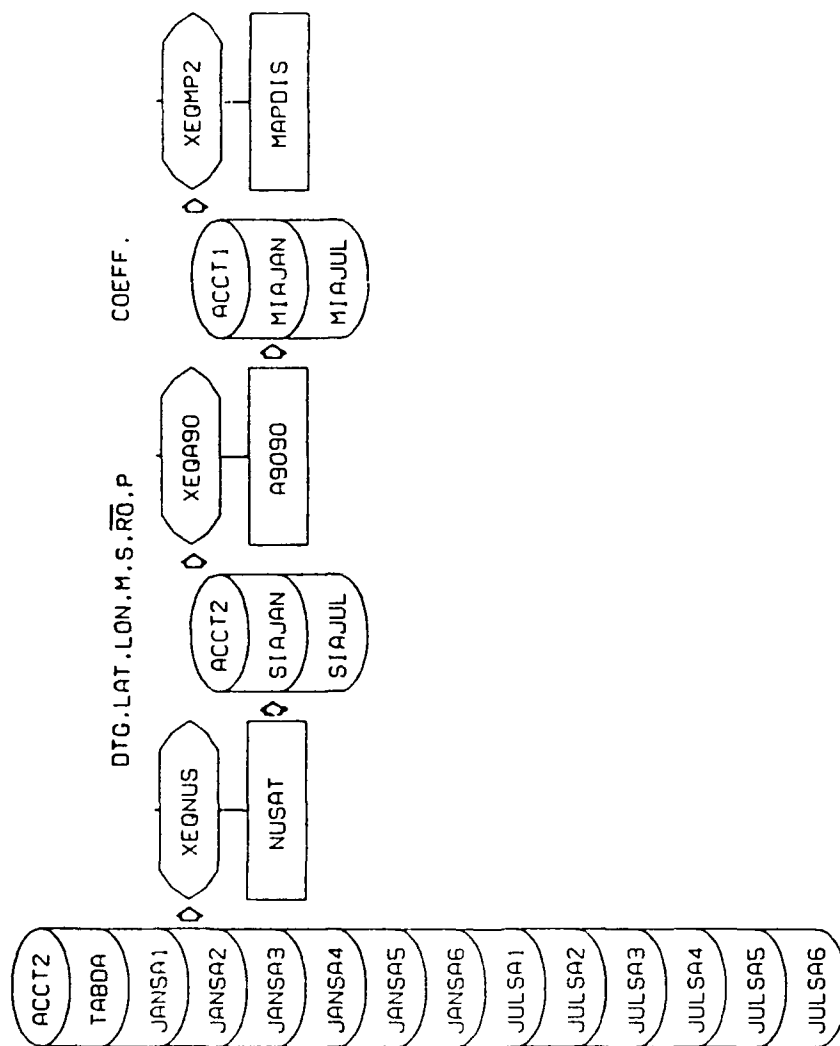


Figure 3.5. Data Flow Configuration for Processing Monthly Interannual Variance of Cloud Cover Using NIMBUS 7 CMatrix Data.

January data for ascending and descending nodes are read and combined primarily to compute the monthly cloud cover for each ERB target area, first for the six Januarys and then for the six Julys. The standard deviations of the mean cloud covers are then computed using the standard method for computing standard deviation which is shown below.

$$XM = \sum_{i=1}^N M_i \quad (3.0)$$

$$XM2 = \sum_{i=1}^N M_i^2 \quad (3.1)$$

N = Population ($1 \leq N \leq 6$)

then

$$STD = \pm \sqrt{\frac{XM2}{N} - \left(\frac{XM}{N}\right)^2} \quad (3.2)$$

where

XM = Monthly mean

N = Population (number of years)

i = Year 1 through 6

STD = Standard deviation

A date time group, latitude-longitude, mean cloud cover, standard deviation, and population is output for each ERB target area in files named SIAJAN for January and SIAJUL for July. A null value is output for the parameter ROBAR. The format of the output resembles the one described in Section 2.1.2.

3.3.2 Program A9090

Program A9090, a version of W9090 with the weighting function suppressed, is used to compress the monthly interannual standard deviations in the form of coefficients. Resulting coefficient data files are named MIAJAN and MIAJUL.

3.3.3 Program MAPDIS

Procedure XEQMP2 is a modified version of XEQMPS shown in Figure 2.1. The modification consisted of a new pattern map input table used for defining pattern map codes versus standard deviation values. The procedure was executed to make use of program MAPDIS to produce pattern maps of monthly interannual standard deviations. Figure 3.6 shows the resulting pattern map of interannual standard deviations from the January CMatrix satellite data.

STD% STD%
1 < 1
1 < 2 < 2
2 < 4 < 4
4 < 6 < 6
6 < 8 < 8
8 < 10 < 10
10 < A < 12
12 < B < 14
14 < C < 16
16 < D < 18
18 < E < 20
20 < F < 25
25 < #

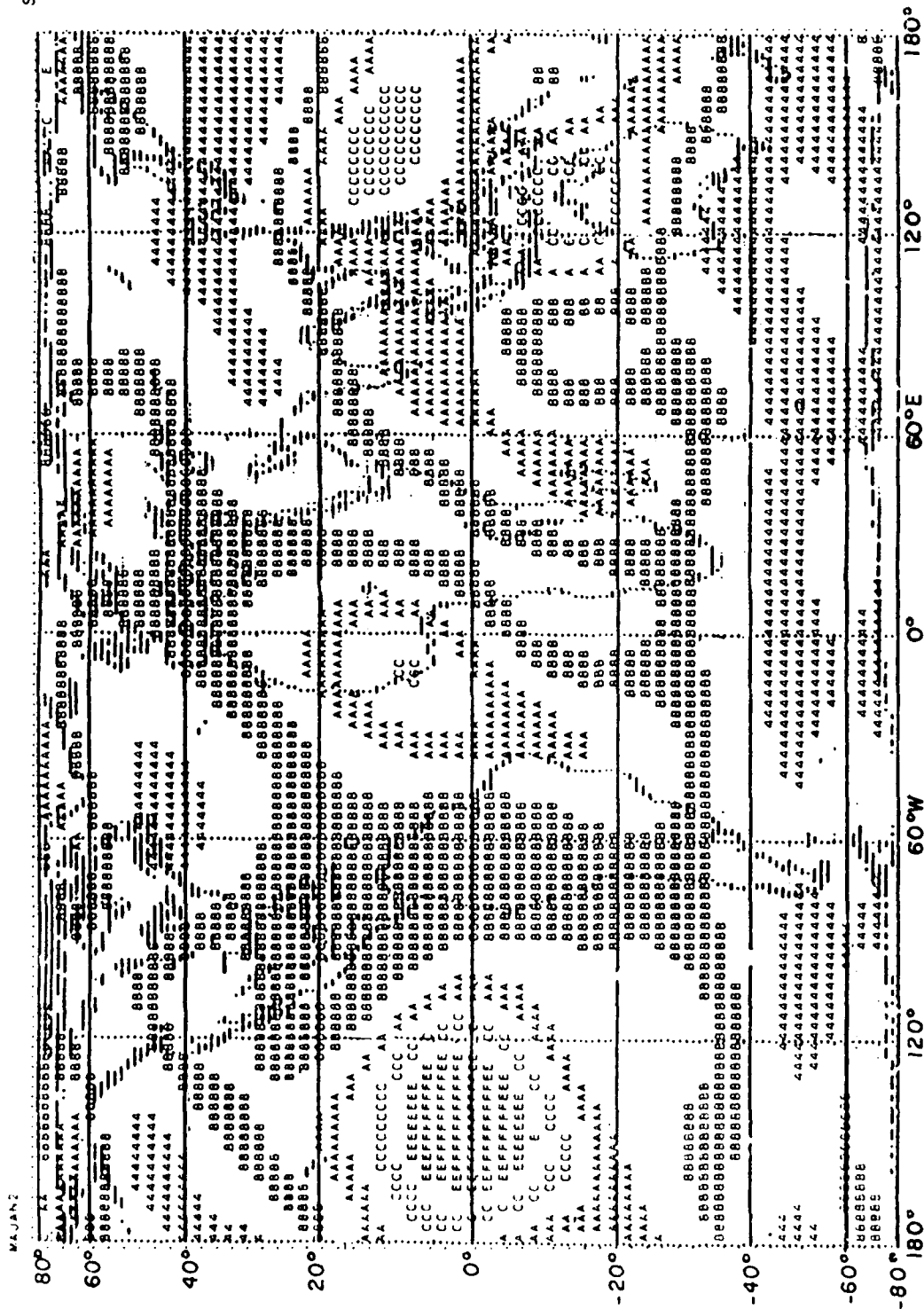


Figure 3.6. Pattern Map of Interannual Standard Deviations of Cloud Cover from Six Januarys Using NIMBUS 7 CMatrix Data.

4.0 BURGER SKY COVER DATABASE DESCRIPTION

The Burger data (Burger, 1985), is a climatology of ground-observed sky cover observations recorded over single-point locations. The archive contains station names, latitude-longitude locators, normalized sky cover distributions, mean sky covers, scale distances, and in some cases, period of record for over 2,000 weather reporting stations around the globe. The data are archived for the four mid-season months of January, April, July, and October where 3 hourly averages of sky cover are centered at the four local standard times of 01, 07, 13, and 19 LST. Data over the ocean areas are archived for the same four mid-season months but sky cover values are monthly means.

The Burger data were compiled mainly from the Revised Uniform Summaries of Surface Weather Observations (RUSSWOs), Naval Intelligence Survey (NIS), and NAVATLAS records. Willand (1988) reformatted the Burger data for subsequent use in the development of an automated global cloud climatology. In fact, this reformatted database was successfully utilized in the development of the prototype automated global cloud climatology (CLOUDZ) by Boehm and Willand (1988). Further preprocessing of this database to include the new parameter mean correlation is necessary and is discussed below.

4.1 Processing of Burger Sky Cover Data

The data flow configuration for processing the Burger data is shown in Figure 4.0. The main objective of the processing is to extract from the Burger data the mean sky cover, standard deviation, and population, and to compute the mean correlation and store the results in a format like that described in Section 2.1.2. Two methods of accomplishing this objective were devised.

4.1.1 Method 1 Program BURFOR

Initially, Method 1 utilizes the revised Burger data files of mean sky cover and scale distance values by Willand (1988) to convert to mean correlation. From Figure 4.0, these files are RUSARC, FINARC, and SEAARC. Program BURFOR utilizes a function SKYRO, listed in Appendix A, to convert the scale distance values found in these files into mean correlation.

4.1.1.1 Scale Distance to Mean Correlation (Subroutine SKYRO)

The computation of mean correlation given sky dome scale distance can be computed by the use of the subroutine SKYRO listed in Appendix A and described below.

(COEFF. LST)

(LST)

OTC.LA.LO.M.SO.RD.P

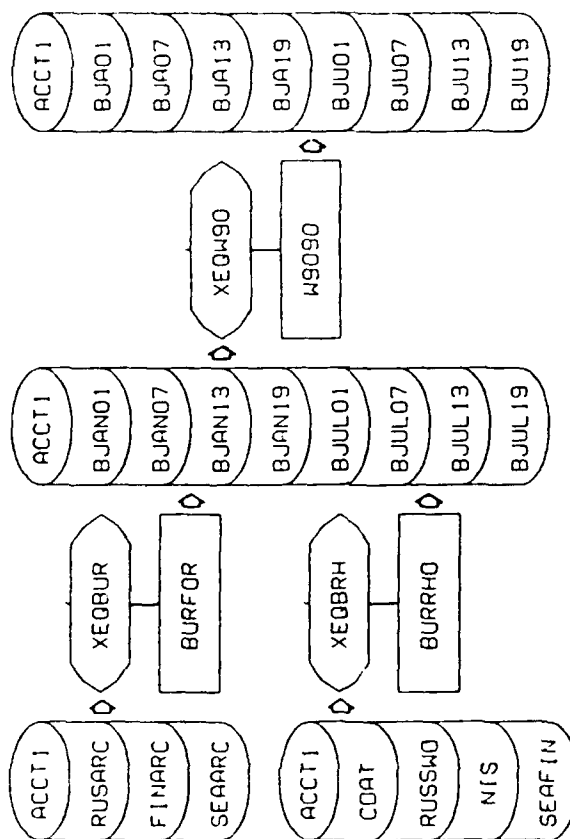


Figure 4.0. Data Flow Configuration for Processing Burger Sky Cover Data into a Format Suitable for Program W9090 Ingest.

CALL Statement: CALL SKYRO(R,RBAR)

Arguments: R = Scale distance value in
kilometers to be converted to
mean correlation

RBAR = Mean correlation returned

Algorithm: Convert scaled distance to mean
correlation. Sum correlation values
computed for all distances between points
within a 21 x 21 grid square area of 2,424
kilometers to compute mean correlation.
The 2,424 km area represents a weather
observer sky dome floor area that has an
assumed radius (r) of 15 nm or 27.78 km.
Thus $\pi r^2 = 2,424$ km circular area.

Step 1. Set N = 21 to define a 21 x 21 array of grid
points within a square.

Step 2. Assuming a 2,424 km sky dome area and a
wavelength of 256, compute the edge of the
square in wavelengths as

$$\text{EDGE} = \frac{\sqrt{2424}}{(256R)} \quad (4.01)$$

Step 3. Compute known distance summations.

$$\text{SUM} = \frac{21 \times 21}{4} \quad (4.02)$$

Step 4. Set delta x as

$$DX = \frac{EDGE}{N-1} \quad (4.03)$$

Step 5. For I = 1 to N-1

For J = 0 to N-1

Compute SIG (Figure 4.0.1)

$$SIG = \sqrt{I^2 + J^2} \cdot DX \quad (4.04)$$

Compute fractional distance as

$$D = SIG - INT(SIG) \quad (4.05)$$

Compute the correlational coefficient
using the 3D-BSW model by Gringorten
(1987)

$$\rho = 1 - 3D + 2D^2 \text{ For } 0 \leq SIG \leq 1 \quad (4.06)$$

or

$$\rho = D \left(\frac{1 - 3D + 2D^2}{SIG} \right) \text{ for } SIG > 1 \quad (4.07)$$

Sum ρ as

$$Sum = Sum + \rho (N-I) (N-J) \quad (4.08)$$

Step 6. Compute mean correlation $\bar{\rho}$ as

$$RBAR = \frac{4SUM}{N^4} \quad (4.09)$$

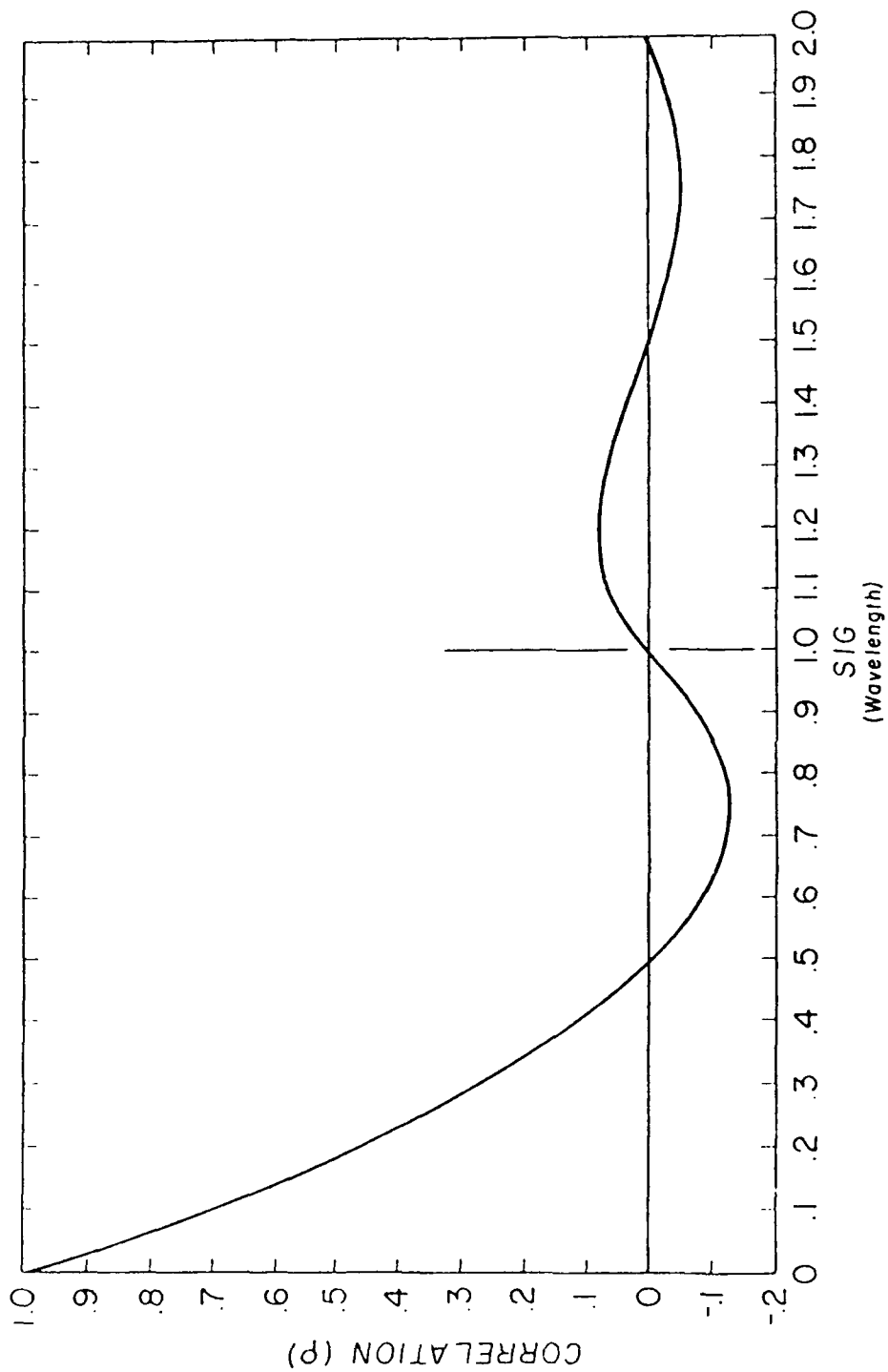


Figure 4.0.1. The Relation of Correlation ρ to Distance SIG in the 3D-BSW Model. The Unit of Length for SIG is the Wavelength.

Subroutine SKYRO is used to convert scale distance values found in the SEAARC and FINARC files to values of mean correlations. The conversion algorithm is particularly useful when applied to the scale distances stored in the FINARC file because 1) the file is quite large and can be processed more efficiently with the use of SKYRO and 2) the original sky cover distributions in the NIS file used by Burger to compute the FINARC mean sky covers and scale distances were archived into only two predetermined sky cover density categories, the probability of sky cover less than or equal to F1 and greater than F2. Therefore, computations of mean sky cover and variance by the methods in Section 4.1.2.1 would likely be inaccurate using only two sky cover categories.

4.1.1.1.1 Testing SKYRO Using Program BURG

A test was performed to compare values of mean correlation computed from the tetrachoric correlation method with those derived using subroutine SKYRO. Predicted sky cover distributions using the Burger technique were also compared with those using mean correlation and the CUB (Cumulative Distribution by Boehm, to be published) technique. These tests were performed using several station data samples from the Burger NIS data archive. The sample tested from the NIS data archived for Amsterdam, Netherlands, in October at 1900 LST is shown in Figure 4.0.2.

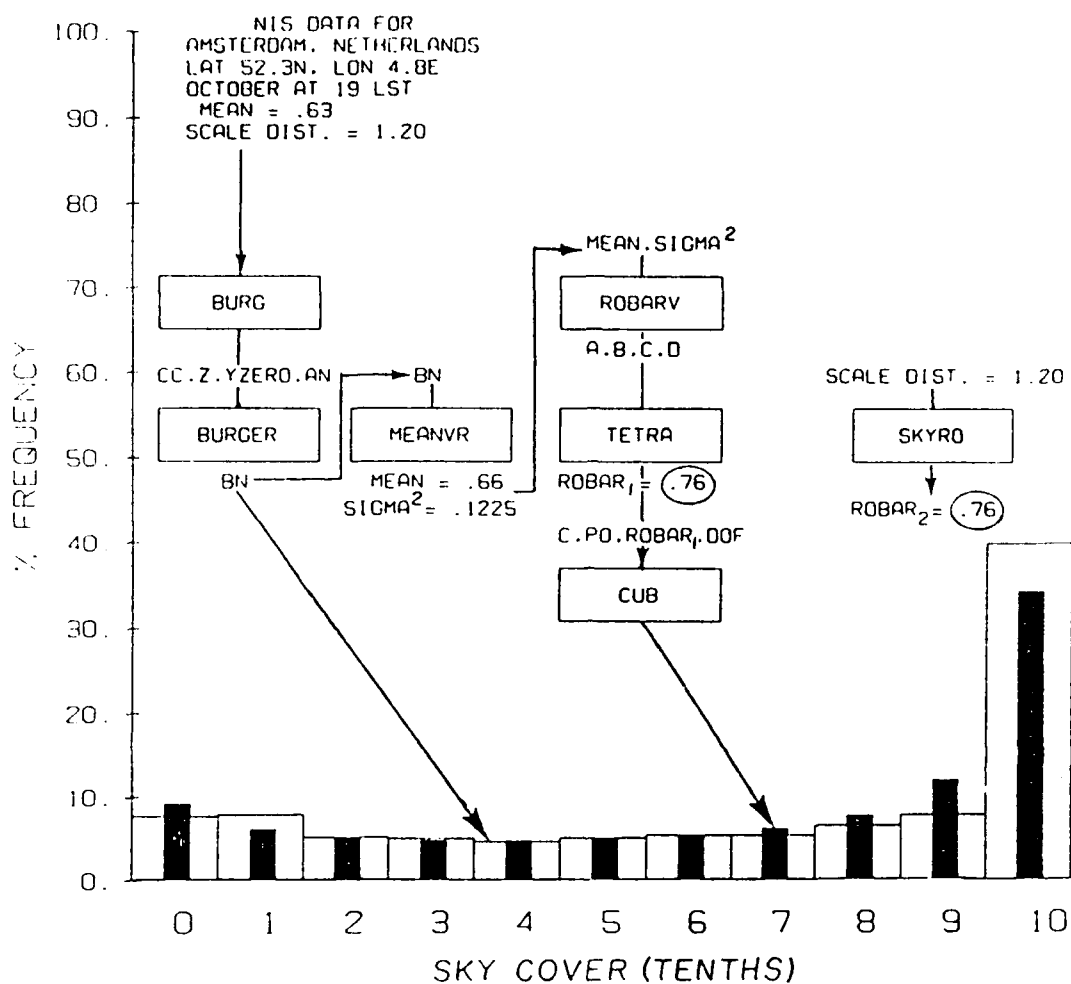


Figure 4.0.2. Comparisons of Mean Correlation ($ROBAR_1$) as Derived Using Mean Sky Cover and Variance from a Predicted Sky Cover Distribution Using the Burger Method (Open Histogram) with Mean Correlation ($ROBAR_2$) Computed Using Routine SKYRO. The Bar-like Histogram is the Distribution Predicted Using CUB.

The data flow for program BURG is superimposed onto histogram plots in Figure 4.0.2 to depict the steps taken to perform comparisons of mean correlation derived from mean sky cover and variance given a Burger derived sky cover distribution with the mean correlation computed by SKYRO.

From the two probability sky cover distribution ($F1 \leq .19$, $F2 > .52$) for Amsterdam, Burger computed a mean sky cover of .63 and a scale distance of 1.20. These values were input to program BURG which calls the routines necessary to perform the testing by the steps outlined below.

Step 1. Routine BURGER is called to generate from a mean sky cover and scale distance a predicted sky cover distribution in tenths.

Inputs to the routine and returned values are
cc = a counter of .0 to .9 in increments of .1

$$Z = \left(\frac{\ln(49.234)}{\frac{\text{scale distance}}{\ln(2.0)}} \right) \quad (4.09.1)$$

YZERO = -ENORM (mean sky cover)

where

ENORM is a FORTRAN function for converting probability into equivalent normal deviate (END). Thus, YZERO is the END of mean sky cover.

AN = Returned cumulative probability of
sky cover for category cc.

The cumulative distribution generated is then converted to a sky cover density distribution (BN) and plotted out in histogram form.

Step 2. This Burger derived distribution (BN) is sent to routine MEANVR that computes from the distribution the mean sky cover and variance (Sigma^2).

Step 3. The mean sky cover and variance is input to routine ROBARV which computes A, B, C, and D for input to the tetrachoric routine TETRA. ROBARV uses the equations 2.2 through 2.5 in Section 2.1.2 to perform the conversion.

Step 4. Routine TETRA using A, B, C, and D from above computes mean correlation (ROBAR).

Step 5. The 12 July 1990 version of CUB is then called to generate the dark bar-like histogram in Figure 4.0.2. for comparisons with the BURGER distribution using the augments listed below.

$c =$ coverage, i.e., the fraction of the domain
 $0 \leq c \leq 1.0.$

$p_0 =$ mean probability for a point in the domain
 $0 < p_0 < 1.$

ROBAR = mean correlation between points in the
domain $0 \leq \text{ROBAR} \leq 1.0.$

DOF = degrees of freedom (80) which is related
to number of pixels used to observe the
domain

Step 6. Finally, routine SKYRO is called with the
scale distance value from the NIS file to
compute independently a mean correlation
(ROBAR_2).

As shown, ROBAR_1 compares exactly with
 ROBAR_2 to two decimal places for Amsterdam.
The CUB distribution compares favorably with
the Burger distribution. The error in the end
categories of the distributions may be due to
observer bias between completely clear sky or
completely overcast.

Method 1 revealed favorable results using data in files FINARC and SEAARC, but it uncovered several undetected errors made previously in sky cover distributions archived in the RUSSWO file of the original Burger database. These events in the database were corrected but it was decided that instead of using mean and scale distance parameters archived in the RUSARC file for computing mean correlation, that more reliable values of mean correlation could be computed using the corrected Burger RUSSWO sky cover distributions for a given station and time. This led to the development of Method 2 which is described next.

4.1.2 Method 2 Program BURRHO

Burger sky cover distributions are found in the data set called CDAT (bottom left side of Figure 4.0). This file is further broken down into three differently formatted files each containing a different type of normalized sky cover distribution. These files are called RUSSWO, NIS, and SEAFIN. Program BURRHO utilizes the sky cover distributions in file RUSSWO to compute the mean sky cover and variance which can then be used with the TETRA function to finally compute mean correlation.

4.1.2.1 RUSSWO Distributions to Mean and Variance for Mean Correlation

Given a sky cover distribution of tenths of sky cover amounts like that archived for RUSSWO data, mean sky cover (\bar{a}) and variance (v) of the distribution can be computed as follows:

$$\bar{a} = \sum_{i=1}^N p_i a_i \quad (4.0)$$

where

p are the unconditional sky cover probabilities for N sky cover amount categories.

a are the mean sky cover amounts in each category where N is 11 in this case (sky cover amounts used are .025, .1, .2, .3, .4, .5, .6, .7, .8, .9, .975).

Standard deviation of the mean sky cover amount is computed as follows:

$$\bar{a}^2 = \sum_{i=1}^N p_i a_i^2 \quad (4.1)$$

where

p and a are as defined above

then

$$\sigma = \pm \sqrt{\bar{a}^2 - (\bar{a})^2} \quad (4.2)$$

finally

$$v = \sigma^2 \quad (4.3)$$

where

σ = standard deviation

v = variance

N = 11

Mean correlation can now be computed using the formulas 2.2, 2.3, 2.4, and 2.5 in section 2.1.1 and the tetachoric correlation function listed in Appendix A and described by Smyth, et al. (1990).

Programs BURFOR and BURROW write a date time group, latitude-longitude, mean sky cover, standard deviation, mean correlation, and population for each Burger station in the same format as that described in Section 2.1.2. These records are stored as file names like BJA01 meaning Burger data for January at 0100 LST.

4.1.3 Program W9090

Program W9090 is executed next to compress the data in the form of coefficients that are stored under file names BJA01 meaning Burger data for January at 0100 LST.

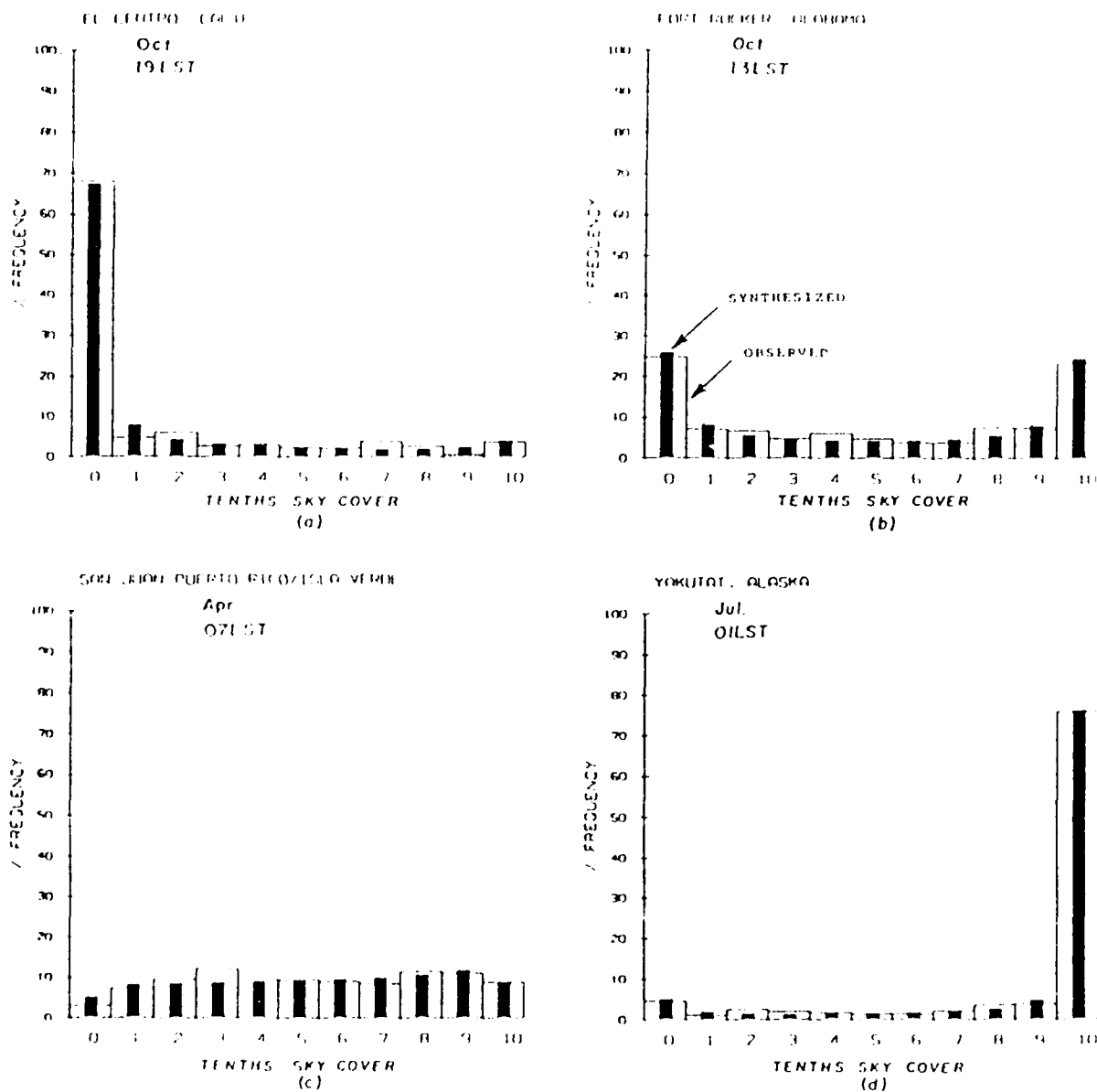


Figure 4.1. Comparison of CUB Synthesized Sky Cover Distributions (Dark Bar) with Burger RUSSWO Data Distributions (Open Histogram). The Cases Shown Represent Typical Sky Cover Distributions over the a) Desert, b) Mid-latitude, c) Tropical, and d) High Latitude Locations.

4.2 Comparisons of Burger RUSSWO Sky Cover Distributions with CUB Predicted Distributions

By utilizing the Burger RUSSWO data, four sets of distributions were plotted in histogram form, Figure 4.1, and the sky cover distributions predicted by the Boehm (CUB) algorithm were superimposed onto the histograms as a dark bar for comparison. The stations were chosen to show results for typical sky cover distributions over (a) desert areas, (b) mid-latitudes, (c) tropicals, and (d) high latitudes.

5.0 DATSAV DATABASE DESCRIPTION

Many hourly surface observations taken at individual stations around the world have been formatted and archived on magnetic tape by the U.S. Air Force Environmental Technical Applications Center (USAFETAC). This archive of surface observations is called DATSAV. The archived hourly data consists of weather conditions such as temperature, winds, dew point, total sky cover, cloud types, etc. Documentation of the original DATSAV database (Ref. Manual, 1977) has been replaced by the updated DATSAV2 documentation (Ref. Manual, 1986). According to the new manual, the updated DATSAV2 format offers two major improvements over the original DATSAV. First, the DATSAV2 format conforms to Federal Information Processing Standards. Second, the DATSAV2 format can accommodate meteorological data fields that could not be handled in older formats.

5.1 Processing of DATSAV and DATSAV2 Data

DATSAV data tapes for 66 stations positioned in clusters at various locations around the world were successfully processed by Willand (1988) for purposes of archiving hourly mean sky cover and scale distance parameters. Since then 26 additional stations within new clusters have been added to augment this data collection. These newly acquired samples were in the updated DATSAV2 format. Investigation of the new

format reveals that no changes are required to the original DATSAV data extraction program, GETCOV (Willand, 1988), which is used for extracting the hourly sky cover values. However, procedure XEQGET, which is used to run GETCOV, had to be changed to process multiple files within a single DATSAV2 tape. Figure 5.0 depicts the general positions of all the acquired DATSAV clusters. The data flow for processing the DATSAV2 and DATSAV data is shown in Figure 5.1.

5.1.1 Program GETCOV

The flow of data processing begins with the newly acquired DATSAV2 data tapes.

Procedure XEQGET positions the DATSAV2 data tape to the proper file so that program GETCOV can extract the hourly sky cover observations from a selected station. These hourly sky cover observations from each selected station are packed and stored on disk in random access binary data files. Detailed documentation of the packing scheme for random access binary data storage of the sky cover observations can be found in Willand (1988). After processing the newly acquired DATSAV2 data through program GETCOV, backup tape M5693 containing DATSAV hourly sky cover observations processed by Willand (1988) were spooled back onto disk storage. Hourly sky cover data for all the selected stations in each cluster were then ready for further processing of two specific tasks, A and B.

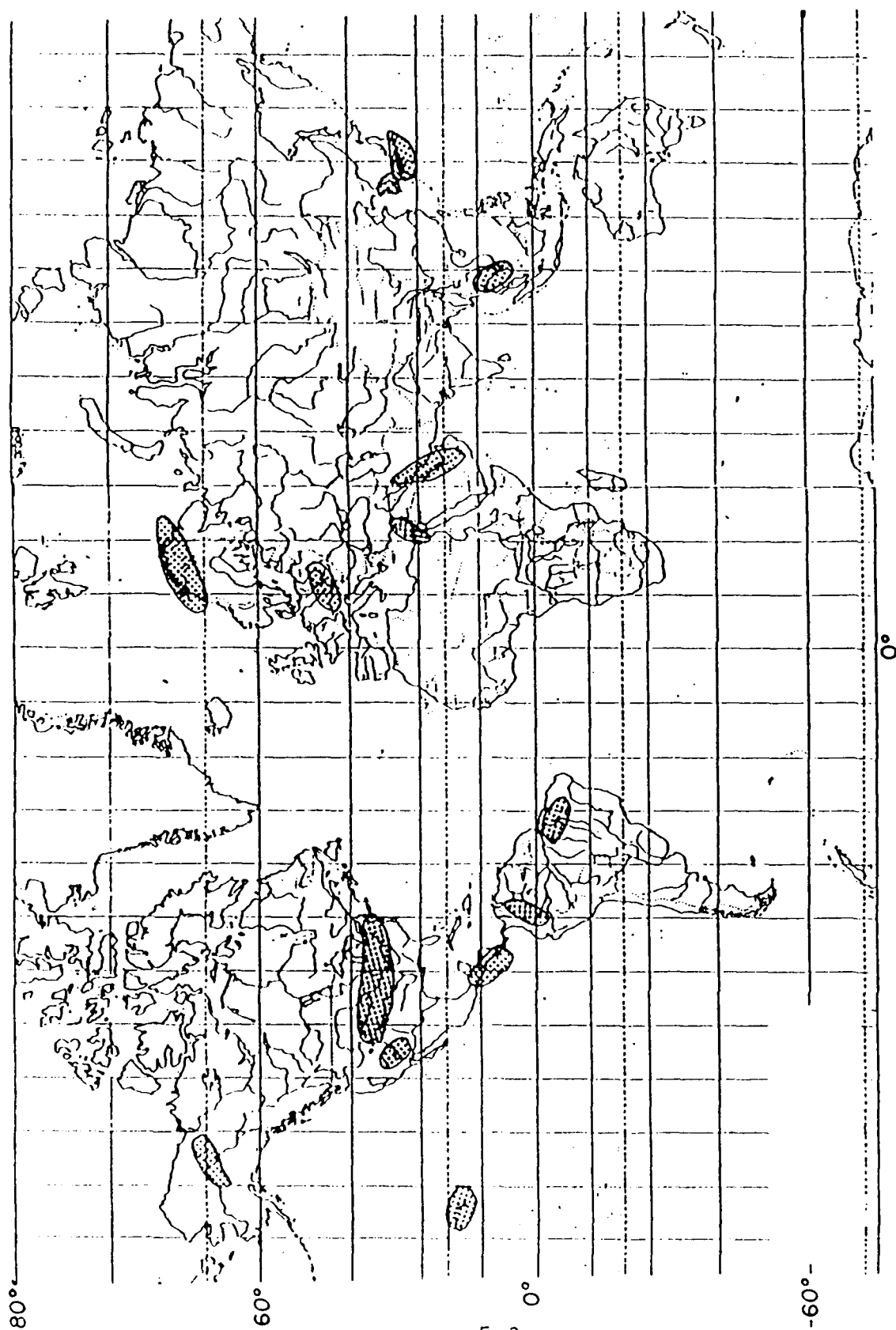
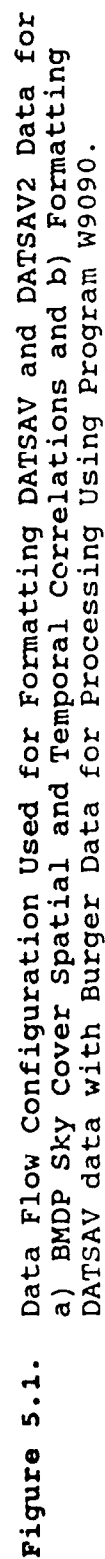


Figure 5.0. Map of Cluster Positions for DATSAV and DATSAV2 Data Collections.



In task A, the DATSAV sky cover data had to be put into a format suitable for use with the Biomedical Data Processing Statistical Software Package (BMDP) (Dixon, et al., 1988) for deriving spatial and temporal correlation of sky cover over stations within separate clusters. In task B, the data had to be formatted so that it could be merged with Burger data for compaction using program W9090. The following sections describe the data flow.

5.1.2 Formatting for Spatial and Temporal Correlation of Sky Cover

A formatting program FORDAT was written to format DATSAV sky cover data into an acceptable form for the BMDP package to process.

5.1.2.1 Program FORDAT

Procedure XEQFDA, used for executing program FORDAT, requires a FORTRAN NAMELIST input file to define input parameters to the program to format the DATSAV sky cover data for BMDP processing. The necessary NAMELIST arguments are as follows:

\$IDEF

PRINT = .TRUE., Print resulting data values

IY1 = 73, First year of data extraction

IY2 = 83, Second year of data extraction
IM1 = 7, First month of data extraction
IM2 = 7, Second month of data extraction
NUMT = 12, Pointer to output file

\$END

The format used by program FORDAT to write the new data files for BMDP processing is as follows.

```
WRITE(NTAPE,100) ISTA,XLAT,XLONG,IY,IM,ID, (IDATA(I,ID) ,  
I=1,IT1,IT2)  
100 FORMAT(I6,F6.2,F7.2,3I2,24I2)
```

where

ISTA = Station Number
XLAT = Station Latitude
XLONG = Station Longitude
IY = Year
IM = Month
ID = Day
IT1 = Start Time (lst)
IT2 = End Time (lst)

The formatted ASCII data files are first output to storage on the CYBER, then transmitted over to the VAX machine for processing spatial and temporal correlations of sky cover using the BMDP package.

5.1.3 Formatting DATSAV Data for W9090 Program Runs

Another formatting program was needed to format the DATSAV sky cover data for use with program W9090.

5.1.3.1 Program FORDA2

Program FORDA2 is a formatting program that is used to format DATSAV sky cover data into a form that can be merged with the Burger database for final processing using program W9090. FORDA2 reads and unpacks the DATSAV packed binary sky cover data, converts GMT to LST, and computes for mid-season months the three hourly mean sky cover values centered at 01, 07, 13, and 19 LST for a given station to match the structure of the Burger data. Variances of sky cover are also computed and routine TETRA is utilized again to compute mean correlation. Finally, the parameter mean sky cover (VM), standard deviation (VS), mean correlation (ROBAR), and population (IXP), are output in the format described below.

```
WRITE(NTAPE,100) ITC, ISEA, ITM, ITIM, XA, XO, VM, VS, ROBAR, IXP
100 FORMAT(13,12,I1,I2,F6.2,F7.2,3F6.3,I8)
```

where

```
ITC   = Data code
ISEA  = Season
ITM   = Time code
```

ITIM = Time (LST)
XA = Latitude
XO = Longitude
VM = Mean sky cover
VS = Standard deviation
ROBAR = Mean correlation
IXP = Population count

These output files named D101 or D107, etc., for station D, month 1 at 01 and 07 LST, etc., are merged with the Burger BJ data files which are then read by program W9090 to derive the coefficients for data compaction. A list of stations within clusters processed to date from DATSAV and DATSAV2 data tapes is shown in Figure 5.2. A plot of the temporal sky cover correlation derived from BMDP processing for stations in the midwestern U.S. cluster for January is shown in Figure 5.3.

Sta.	Lat.	Lon.	Elev.	Name	Sta.	Lat.	Lon.	Elev.	Name
MIDWEST:					SOUTH AMERICA:				
72638	44.37	-84.68	0351	HOUGHTON LAKE MI	90370	0.87	-77.68	2975	IPIALES
72652	43.07	-95.53	0423	PICKSTOWNE S.D.	80110	6.22	-75.60	1498	MEDELLIN
72424	37.90	-85.97	0233	FT. KNOX KY	80089	7.00	-74.72	0610	AMALFI ANTIOQUIS
72551	40.85	-96.75	0362	LINCOLN NE	80259	3.55	-76.38	0969	CALI/PARAMESECA
72640	42.95	-87.90	0220	MILWAUKEE	80308	2.45	-76.60	1730	POPA VAN/LEON-VALE
72641	43.13	-89.33	0267	MADISON	AMAZON BASIN:				
72637	42.97	-89.73	0238	FLINT	82678	-6.17	-42.58		FLORIANO/CANGAPARA
72552	40.97	-98.32	0563	GRAND ISLE NE	82562	-5	-49.07		MARABA
72556	41.98	-97.43	0479	NORFOLK NE	HAWAII:				
72423	38.18	-85.73	0151	LOUISVILLE KY	91186	21.15	-157.10	0138	MOLOKAI
72635	42.80	-85.52	0242	GRAND RAPIDS MICH	91182	21.35	-157.93	0005	HONOLULU
72636	43.17	-86.23	0191	MUSKEGON MI	SWEDEN-NORWAY:				
72529	43.12	-77.67	0171	ROCHESTER NY	01006	69.07	15.18	0014	ANDA
72527	42.93	-78.73	0217	BUFFALO NY	01025	69.68	18.92	0010	TRONISO
72545	41.88	-91.70	0263	CEDAR RAPIDS IOWA	01059	70.07	24.98	0008	BANAK
72546	41.53	-93.67	0292	DES MOINES IOWA	01041	69.75	21.03	0004	NOROREISA
72734	46.47	-84.37	0220	SALT ST. MARIE	10570	69.37	24.43	0285	CUOUTATHAUKI
SOUTHWEST:					10490	69.98	23.35	0003	LUFTHAUN
725690	42.55	-106.28		CASPER WY	01098	70.37	31.10	0014	VARDO
746120	35.41	-117.41	0681	CHINA LAKE CAL	20960	67.20	23.42	0168	PAJALA
723865	36.25	-115.03	0570	NELLIS AFB	22101	68.90	28.45	0118	VIRTANIFALLI
724660	38.82	-104.72	1881	COLORADO SPRINGS CO	28360	67.37	26.65	0178	SODANKYLA
722800	32.40	-114.36	0063	YUMA ARIZ	22106	68.60	31.80	0034	PADUN
726940	44.92	-123.02	0061	SALEM/ MCNARY	22119	68.32	33.33	0130	PULOZERN
MIDDLE EAST I:					22204	67.55	30.43	0221	KORDOR
400800	33.31	36.18		DAMASCUS, SYRIA	22205	67.60	31.17	0160	ENA
401840	31.46	35.14		JERUSALEM, ISRAEL	22216	67.55	33.39	0136	APATITY
623660	30.00	31.17		CAIRO, EGYPT	22232	67.70	37.12	0244	KOLMYARR
MIDDLE EAST II:					22249	67.13	39.67	0149	KANEUKA
404160	26.18	50.08		DHARAN, S. ARABIA	GERMANY:				
404380	24.31	46.47		RIYADH, S. ARABIA	10147	53.63	9.98	0016	HAMBURG
405820	29.04	47.59		KUWAIT, KUWAIT	67500	46.65	8.62	2287	GUETSCH
406890	30.30	47.47		BASRAH, IRAQ	67620	46.17	8.88	0198	LACARNO
ASIA I:					67700	46.00	8.97	0276	LOGIANO
471220	37.11	127.04	0041	OSAN A.B., S. KOREA	10235	53.06	9.83	0077	SOLTAU
470580	39.03	125.48		PYONG YANG, N. KOREA	10803	48.02	7.85	0300	FREIBURG
ASIA II:					10908	47.87	8.00	1493	FEILDBERG
476420	35.45	139.21		YOKOTA, JAPAN	16066	45.62	8.73	0211	MILAND
478270	31.24	130.39		KAGOSHIMA, JAPAN	16072	45.87	9.07	1319	MONTE BISINO
ASIA III:					16080	45.43	9.28	0107	MILANO LINATE
484560	13.50	100.29	0004	BANGKOK, THAILAND	06700	46.25	6.13	0436	GENEU/COINTRIN
488200	21.04	105.50		HANOI, VIETNAM	06670	47.48	8.53	0432	ZURITCH
488550	16.08	108.22		DA NANG, VIETNAM	06680	47.25	9.35	2500	SAENTIS
489000	10.46	106.34		HO CHI MINH, VIETNAM	06720	46.22	7.33	0482	SION
983270	15.11	120.32		CLARK A.F.B., PHILIPPINES	06760	46.17	8.78	0379	LOCARNO
466960				TAPEI, TAIWAN	06610	46.82	6.95	0489	PAYERN
CENTRAL AMERICA:					06990	47.13	9.53	0463	VADUZ
786410	14.37	-90.32		GUATEMALA CITY, GUAT.	ALASKA:				
786630	13.45	-89.11		SAN SALVADOR, EL SAL.	70174	66.92	-151.52	0205	BETTLES
787200	14.08	-87.15		TEGUCIGALPA, HON.	70063	70.15	-148.20		PRUDHOE
787410	12.10	-86.16		MANAGUA, NICARAGUA	70265	64.40	-147.7	0128	EIELSON
787600	10.00	-85.00		PUNTA ARENAS, C.R.					
788060	8.92	-79.60		HOWARD A.F.B., PANAMA					

Elev. = Meters

Figure 5.2. Listing of DATSAV Station Clusters.

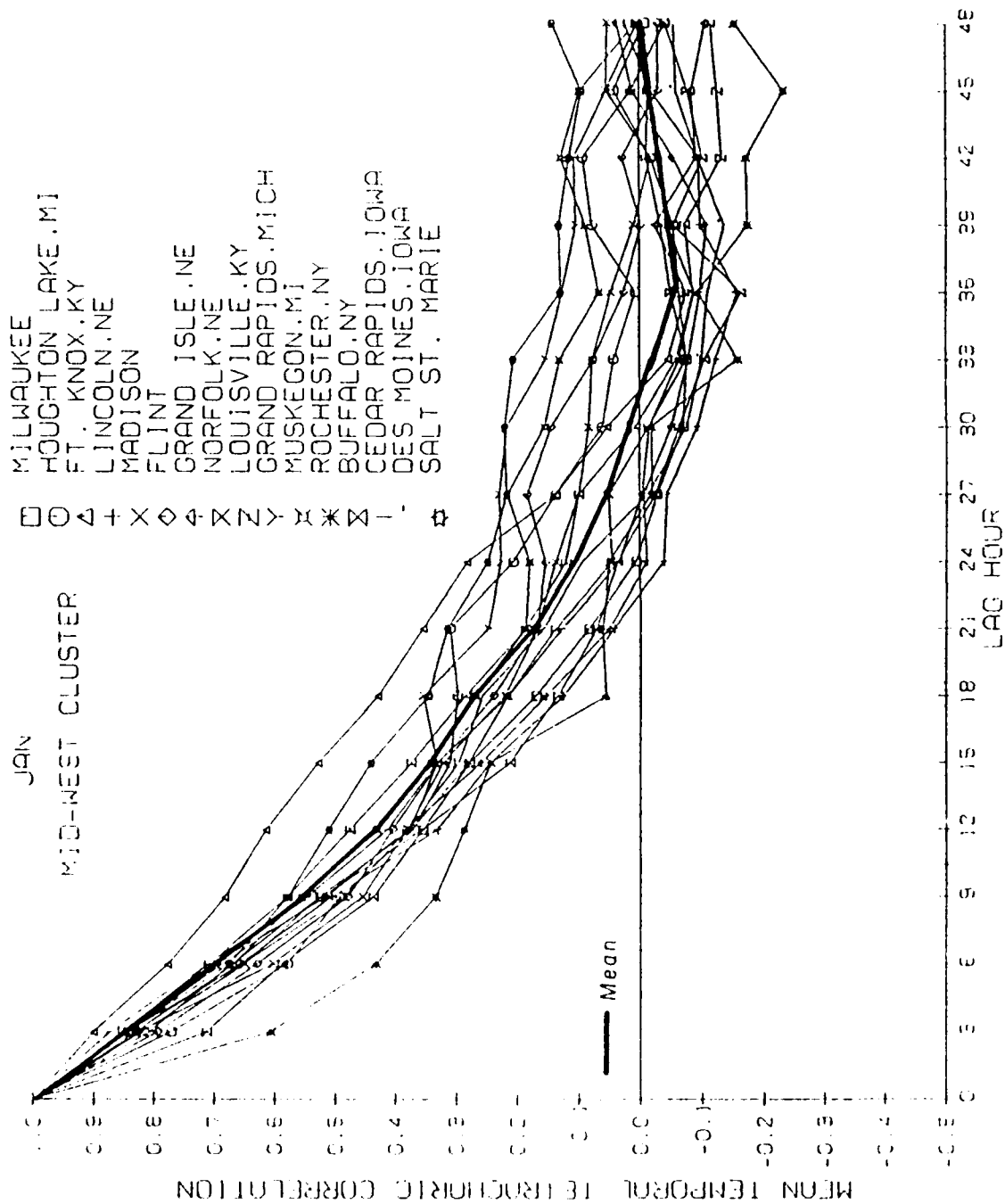


Figure 5.3. Temporal Correlations of Sky Cover for the Midwestern U.S. Cluster.

6.0 SUMMARY OF CLOUD DATABASES

Spatial and temporal attributes of the cloud databases being processed into a climatology of cloud statistics for automated global cloud climatologies are summarized in Table 6.0. Observation of the attributes in this table reveal that none of the databases are alike. For example, each database has its own grid system. DOE data are mapped into a 5 x 5 degree grid system whereas NIMBUS 7 CMatrix data are in a 2070 ERB target equal area grid, but Burger data together with DATSAV data are sky dome point climatologies gridded randomly.

The temporal attributes show that DOE data are stratified by seasons at every three hours GMT while NIMBUS 7 CMatrix data are archived daily at noon and midnight MAST. Burger data, on the other hand, are mid-seasonal monthly means of three hourly mean sky covers centered over four local standard times of day, but DATSAV data are mostly archived 24 hours daily in GMT. These differences are of concern when developing a final climatology of cloud statistics that is constructed by combining or blending all the cloud databases into one. The following sections discuss some of the problems associated with database blending together with present and future progress toward a cloud database blender.

TABLE 6.0

Summary Table of Cloud Databases for C Cloud S 1.0

Cloud Data Base	Source	Spatial Attributes			Temporal Attributes		
		Area Resolution	Area	Points per Map	Yearly Strat.	Times of Day	POR
DOE/NCAR	Mean of Sfc Obs in Area	5 X 5 Deg 555Km X 555Km	Near Equal Area	1820	4 SNS	3 Hrly 0,3,... GMT	Land 11 Yrs. 1971-1981 Ocean 54 Yrs. 1930-1983
NIMBUS Sever CMat.	Sun Sync. Satellite	500Km X 500Km	Equal Area	2070	Daily	2 Hrs 00-12 MAST	6 Yrs. 1979-1985
BURGER	Sfc Obs at a Pt.	Sky Dome Area	Point Clim.	2300	Mid SNS Monthly	LAND 4 Hrs 1,7 13,19 LST OCEAN Monthly 00 LST	24 Yrs. 1945-1973 122 Yrs. 1854-1975
DATSAV	Sfc Obs at a Pt.	Sky Dome Area	Point Clim.	92	Hourly	24 Hrs GMT	11 Yrs. 1973-1983 and 1977-1986

6.1 Database Blending

Further processing of the cloud databases will be necessary in order to adjust the spatial and temporal attributes to a common dimension in both space and time prior to any data blending. Tentatively, the additional processing activities will require interpolators that perform the tasks listed below.

1. Adjustment of cloud databases to an equally spaced grid system
2. Interpolation of cloud parameters to any month, day, and time of day
3. Interpolation or extrapolation of values to adjusted equal viewing area sizes
4. Incorporation and interpolation of terrain induced cloud effects

The problem of adjusting the databases to an equally spaced grid system will most likely be handled by the Fourier-Legendre spatial coefficients derived and processed through the methods described in Sections 2 through 5 of this report.

Interpolation of spatial coefficients for predicting mean sky cover, mean correlation, and population to any month, day, and time of day can be accomplished using Fourier analysis. Since cloud amounts over the globe are influenced by the sun, it would be desirable to design the Fourier analysis system to convert all cloud databases to mean apparent sun time. A partial version of this Fourier system to interpolate values to any time of day was designed and FORTRAN coded into a program called MAST. An exercise was then initiated to test the use of the Fourier system to convert DOE spatial coefficients stratified in GMT to temporal coefficients. The temporal coefficients in turn were then used to create a new set of spatial coefficients that predicted cloud parameters at any time of day in MAST. The final objective of the exercise was to utilize the newly derived spatial coefficients to construct a three-hourly global atlas of mean sky cover, mean correlation, and population in mean apparent sun time for both winter and summer seasons. The resulting atlas was then scrutinized to ensure that the Fourier analysis was performing properly.

6.1.1 Interpolating Data in GMT to Any Time of Day in MAST

Figure 6.0 depicts the data flow for interpolating GMT spatial coefficients derived from DOE data to temporal coefficients using a Fourier analysis that in turn computes spatial coefficients for predicting parameters at any time of day in MAST. As shown in Figure 6.0, the spatial coefficients

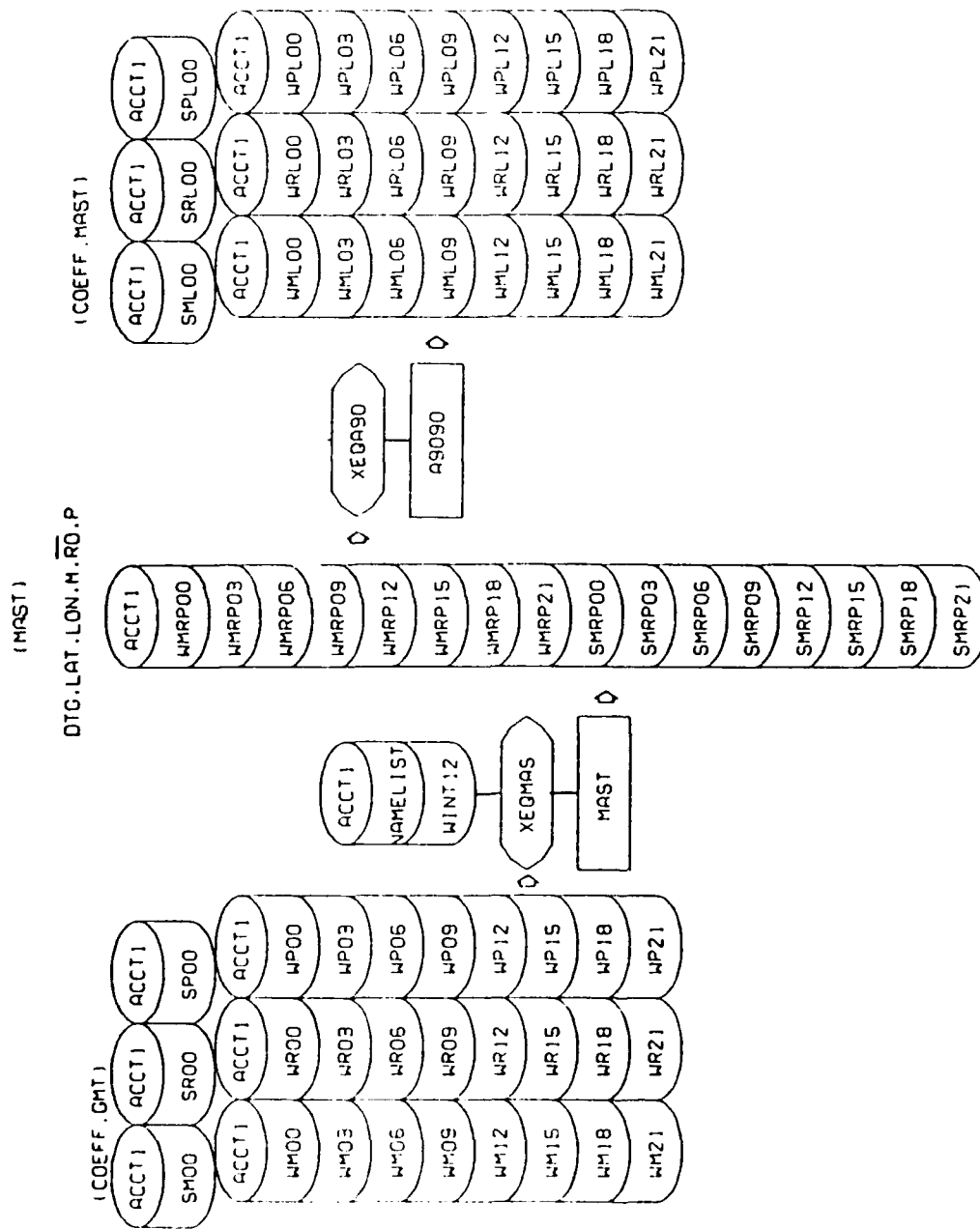


Figure 6.0. Data Flow for Interpolating GMT Spatial Coefficients Derived from DOE Data to Spatial Coefficients for Predicting Parameters in MAST Using a Fourier Analysis.

of mean sky cover, mean correlation, and ln of population for winter and summer at eight GMTs are stored in files with names like WM00, WR00, WP00, meaning winter sky cover (M), mean correlation (R), and population (P) respectively for 00 GMT. These coefficients were derived unweighted from a Fourier-Legendre array size of 10 x 10 or 121 coefficients per data file. (The 10 x 10 array size was used instead of the 14 x 14 array size in order to speed up the computer processing.) The procedure XEQMAS was implemented to execute program MAST which contained the Fourier analysis necessary for performing the interpolation of the spatial coefficients stored in the data files to temporal coefficients.

6.1.2 Program MAST

Program MAST read in a NAMELIST parameter file that defined the desired seasons, MASTs, and names of the data files containing the GMT spatial coefficients that were to be converted to temporal coefficients. The mapping resolution chosen for the exercise being performed was the 5 x 5 degree DOE grid system. Therefore, the DOE data file for winter at 1200 GMT, WINT12, was made available to the program for input of latitude-longitude locators that were necessary for defining and locating points in the DOE grid system. Next, the 121 spatial coefficients for all eight GMTs and for all three parameters were read into an array dimensioned (8 x 121 x 3). The Fourier analysis then performed the

summations on this array to produce the temporal coefficients at GMT that were stored into a second array dimensioned the same as above. These temporal coefficients were then utilized to compute a new set of 121 spatial coefficients for predicting the parameters of interest over any given longitude location at any given MAST. For any given longitude position and MAST, the conversion to GMT for use with the GMT temporal coefficients was made using the following.

$$\text{GMT} = -.06667 \times \text{LONG} + \text{MAST} \quad (6.0)$$

where

LONG = Longitude (Degrees East)

MAST = Mean apparent sun time

GMT = Greenwich Mean Time

Tests were then made to ensure that GMT remained within a 24-hour cycle.

This process was repeated for every point in a DOE grid system for each parameter creating a synthesized data file at every three hours MAST. These data files consisted of a date time group (DTG), latitude-longitude reference, mean sky cover (M), mean correlation (R), and population (P) for winter and summer in a format like that described in Section 2.1.2. Names such as WMRP00 meaning winter, mean sky cover, mean correlation, and population for 00 MAST were assigned to the synthesized data files.

6.1.3 Program A9090

Program A9090 is a version of program W9090 with the weighting function removed. This program was called by procedure XEQA90 to produce again 121 coefficients for predicting a desired parameter value anywhere on the globe for each of the eight given MASTs.

Program MAPDIS, not shown in Figure 6.0, utilized the newly derived spatial coefficients to produce an atlas of the three parameters in the form of pattern maps. Figure 6.1 shows samples of winter pattern maps generated for mean sky cover at a) 00 MAST and b) 12 MAST. Note that sky cover amounts at 12 noon are in general higher than the amount globally at midnight. This is particularly noticeable over southwestern Asia, Indonesia, the Congo Basin, the Amazon, and the Caribbean areas.

Although the exercise discussed above demonstrated the use of the Fourier to interpolate spatial coefficients in GMT to those at any time of day in MAST, interpolating to any month and day remains to be demonstrated. It is anticipated that this will be solved when spatial coefficients for all months and seasons for all the cloud databases become available. The methods described in a recent note on

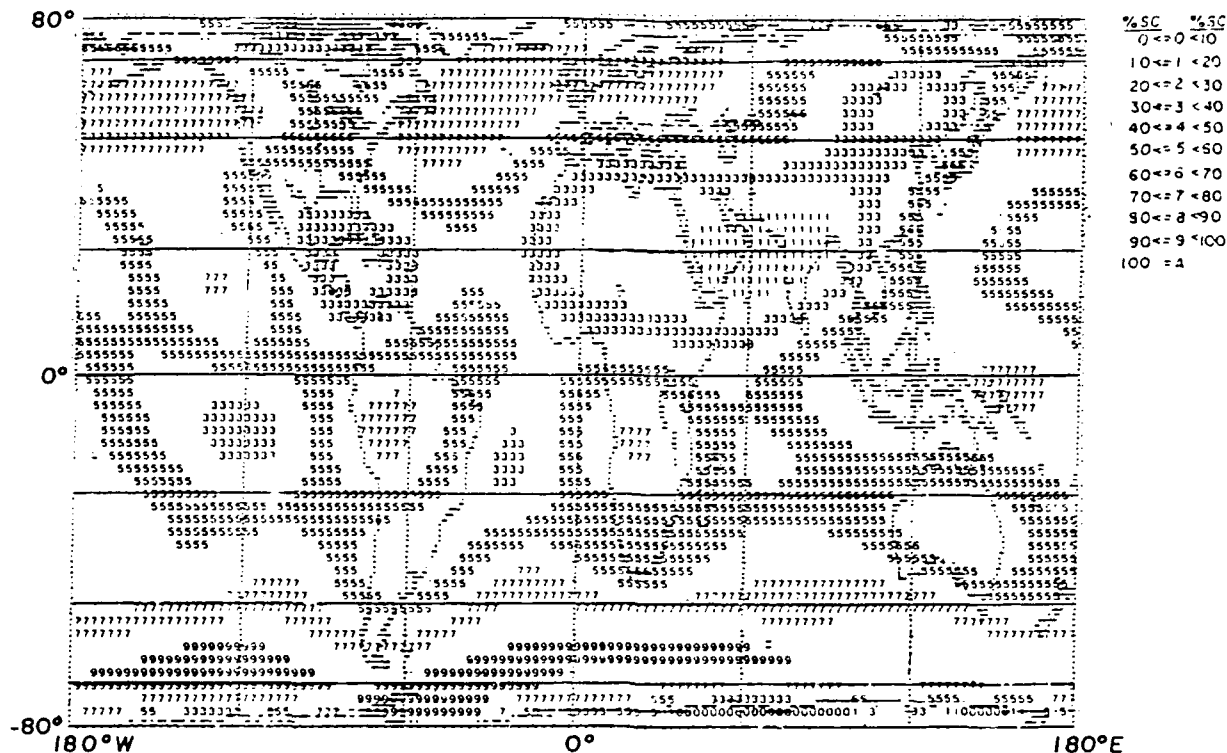


Figure 6.1a. Pattern Map of Predicted DOE Mean Sky Cover (SC) for Winter at 00 MAST.

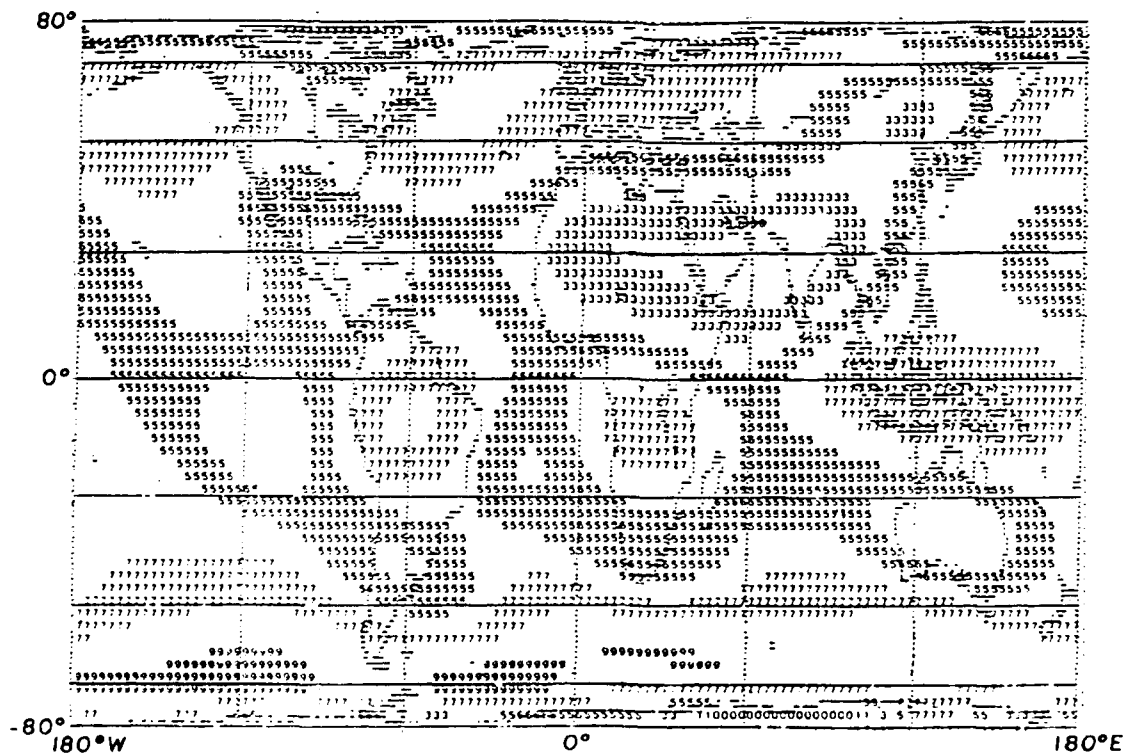


Figure 6.1b. Pattern Map of Predicted DOE Mean Sky Cover (SC) for Winter at 12 MAST.

obtaining daily climatological values from monthly means by Epstein (1991) will be considered as a means of interpolating cloud parameters to adjusted time periods.

6.2 Future Database Blender

Figure 6.2 shows a proposed generalized data flow of preprocessing steps toward future database blending. The blended C Cloud S can be utilized to generate coefficients for implementation of a final automated global cloud climatology.

Sections A and B of Figure 6.2 reflect the processing discussed in Sections 2.0 through 5.0 of this report. In summary, these steps will involve extraction of source data from the DOE, CMatrix, Burger, and DATSAV cloud databases for all seasons and/or months. The extracted data, stored under Section A, will consist of mean sky cover (Mean), mean correlation (ROBAR), ln of population (LN POP), and where possible, monthly interannual variance of sky cover (IA). Spatial coefficients using the Fourier-Legendre analysis in program W9090 will then be produced (Section B of Figure 6-2) for all seasons and months for the data sets archived periods of time. Next, three versions of program MAST may be developed which contain all the interpolation algorithms for adjusting the data sets to a common dimension in space and time. Program BLEND may then be initiated to sum all the

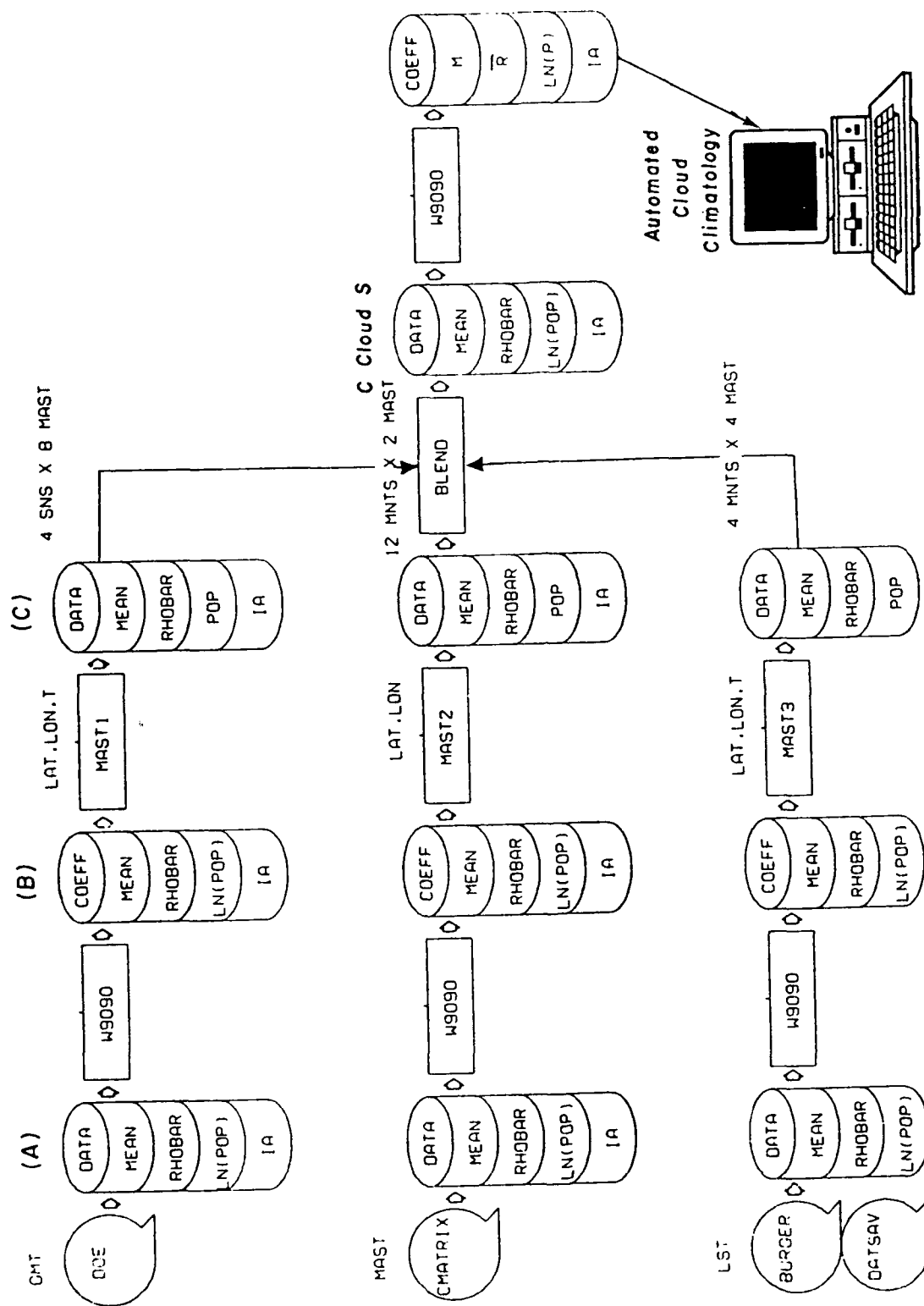


Figure 6.2. Generalized Data Flow Showing Preprocessing Steps Toward Future Database Blending.

synthesized data files shown under Section C into a single C Cloud S utilizing the summation formula in 6.1.

For each point in a data set and for each time period and parameter, compute a final value V by

$$V = \frac{W_1 D_1 + W_2 D_2 \cdots W_i D_i}{\sum_{i=1}^N W_i} \quad (6.1)$$

where

W = Weight (Population)

D = Data value of given parameter

i = Data set

N = Total number of data sets.

Final coefficients can then be transmitted from the main frame to a PC for implementation of an automated global cloud climatology.

7.0 WHOLE SKY PHOTOGRAPHS DATA DESCRIPTION

Whole Sky Photos (WSPs) were collected at the National Weather observing site in Columbia, Missouri, over a three-year period of record from 1966 through 1969 at the three local standard hours of 0900, 1200, and 1500. Figure 7.0a shows an example of a WSP. The sky photographs were processed and utilized in a study by Lund, Grantham, and Davis (1980) for estimating probabilities of cloud-free fields-of-view from the earth through the atmosphere. According to that study, each photograph was processed into a slide which was projected onto a grid (Figure 7.0b). The grid was divided into concentric rings drawn at every five degrees, up to 85 degrees, and the entire circular area was further divided into 185 annual sectors. An estimate was made from each photograph to determine the eighths of cloudiness in each sector. For example, cover amounts tabulated for the WSPs taken on March 3 at noon, 1966, are portrayed in Figure 7.0c. Data estimates were tabulated in a form described next and made available for subsequent computer processing.

7.1 Whole Sky Photograph Data Format

Initial description of the WSP data storage format was described by Lund (1977). This section attempts to update this initial WSP format description.

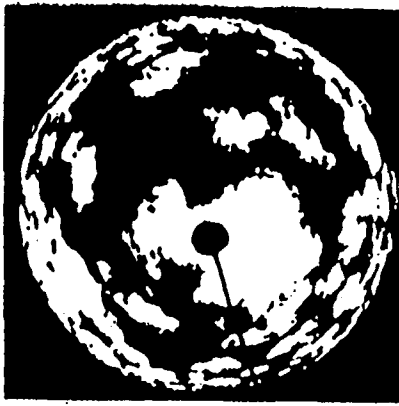


Figure 7.0(a). Example of Whole Sky Photo.

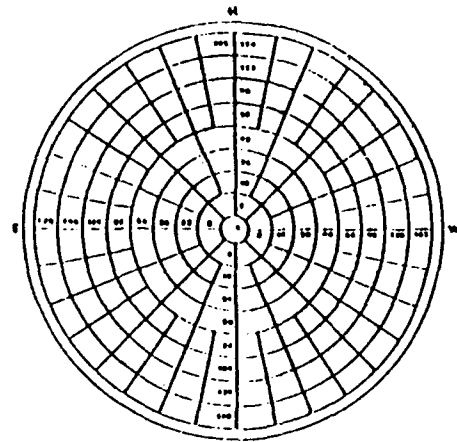


Figure 7.0(b). Whole Sky Photo Grid Used for Reading Sector Sky Cover Amounts.

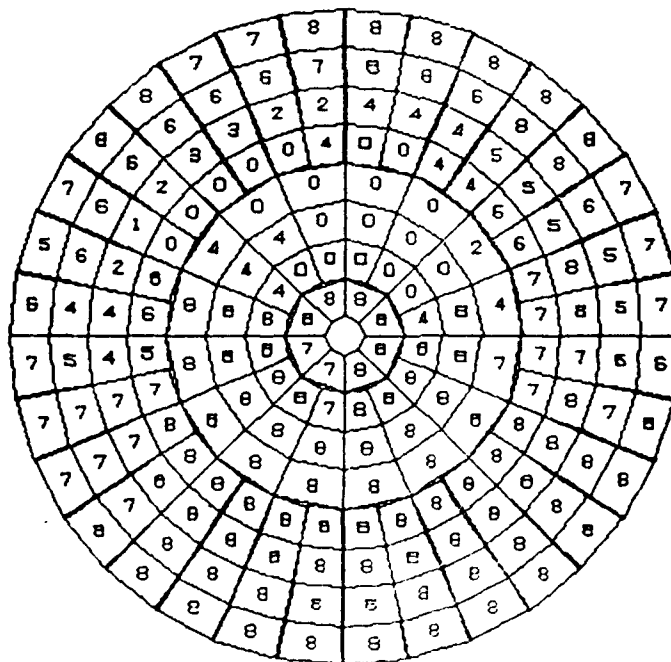


Figure 7.0(c). Sample of Sector Sky Cover Estimates from the Whole Sky Photo Taken at Noon on March 3, 1966.

Back in 1977, tabulated estimates of WSP sky cover were punched onto cards. In addition, data cards containing weather observations from Columbia, Missouri, were obtained that coincide in time with WSP observations. The WSP data ensemble was then assembled sequentially by time in blocks of four cards per observation. The first card describes the weather conditions at the time of the Whole Sky Photograph. The next three cards are the tabulated estimates of sky cover amounts in the sectors over the sky dome shown in Figure 7.0b. A description of the format for reading card one, the weather card, is as follows.

<u>FORTRAN Format</u>		<u>Variable</u>
COL		
1-2	I2	Year
3-4	I2	Month
5-6	I2	Day
7-8	I2	Hour
9	I1	Sky and ceiling
10-12	I3	Surface visibility
13-23	I11	Weather and obstruction to vision
24-26	I3	Total sky cover
27-29	I3	Lowest layer amount
30-31	I2	Lowest layer type
32-34	I3	Lowest layer height
35-37	I3	Second layer amount
38-39	I2	Second layer type

40-42	I3	Second layer height
43-45	I3	Summation total
46-48	I3	Third layer amount
49-50	I2	Third layer type
51-53	I3	Third layer height
54-56	I3	Summation total
57-59	I3	Fourth layer amount
60-61	I2	Fourth layer type
62-64	I3	Fourth layer height
65-67	I3	Total opaque sky cover

This format is in conjunction with the TDF14 tape code AWS NQ form 0-16. The format of the next three cards that contain the sky cover estimates for the 185 sectors are described as follows.

Card 1

Column

1-2	I2	Year
3-4	I2	Month
5-6	I2	Day
7-8	I2	Hour
9	I1	Card #1
10-66	57I1	WSP sky cover estimates in eighths for sectors 1-57
67-71		Blank

Card 2

Column

1-2	I2	Year
3-4	I2	Month
5-6	I2	Day
7-8	I2	Hour
9	I1	Card #2
10-73	64I1	Sky cover estimates for sectors 58-121

Card 3

Column

1-2	I2	Year
3-4	I2	Month
5-6	I2	Day
7-8	I2	Hour
9	I1	Card #3
10-73	63I1	Sky cover estimates for sectors 122-185

Note 1: An asterisk is placed in column 71 of card #1 and column 76 of card #2 and card #3 if some of the whole sky observations were completely clear or completely overcast.

Note 2: For completely cloudy observations, cards number 2 and 3 are reversed in order. This makes no difference since the data on both cards are identical except for

the sequence numbers in column 9. Caution should therefore be exercised in searching by sequence numbers.

The sum of 2,839 hours of observations were accumulated and stored in card image format onto magnetic tape that is discussed next.

7.2 Processing of Whole Sky Photo Data

Several studies have been initiated in the past through processing of the WSP data ensemble described here (Lund, et al., 1972 and Lund, et al., 1980). The following two sections describe 1) newly updated tape archive of WSP data and 2) the processing of WSP data to define the correlation of sky cover over a sky dome. The data flow for processing the WSP data is shown in Figure 7.1.

7.2.1 WSP Tape Archive Update

WSP estimated sky cover data were originally written to magnetic tape M9782 in binary form blocked with certain control words that are not compatible with present day storage schemes. Therefore, it was decided to reformat the data into a form more suitable for present day processing.

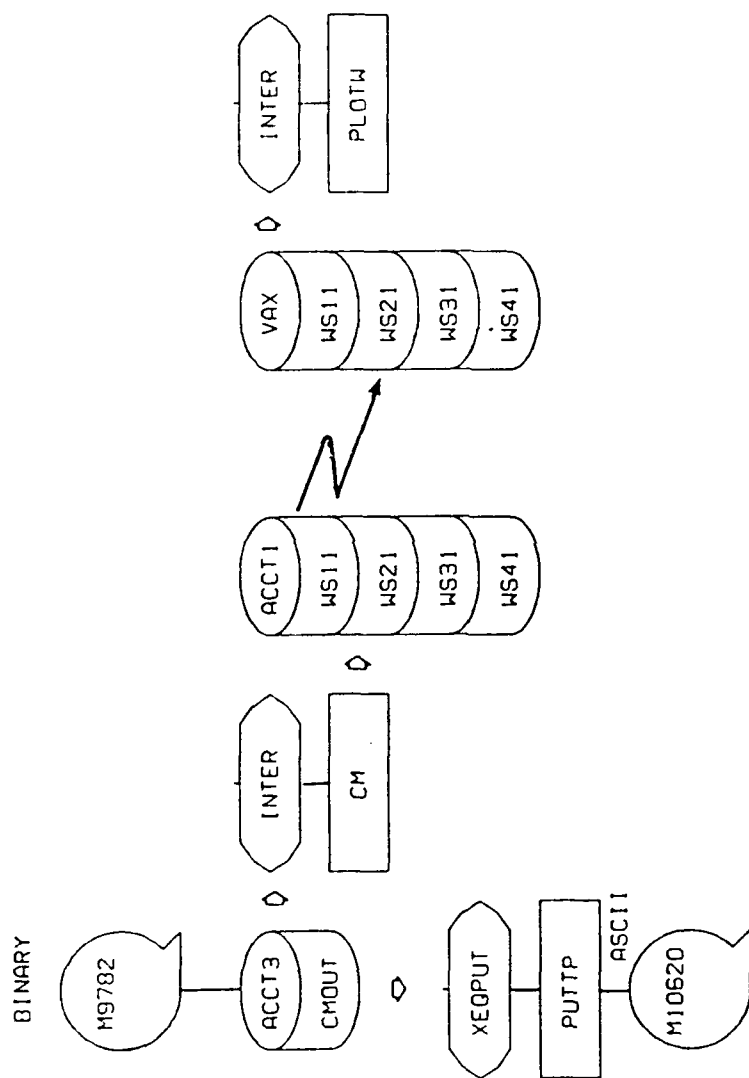


Figure 7.1. Data Flow Configuration of the Processing of Whole Sky Photo (WSP) Data for Tape Storage Update and Correlation Structure of Clouds over a Sky Dome.

7.2.1.1 Program PUTTP

The original tape M9782 was first spooled into a disk storage file named CMOUT. The data from this file were then sent through program PUTTP, called by procedure XEQPUT for purposes of rewriting the data into ASCII card image format. The resulting updated WSP data archive is stored on tape M10620.

The new blocksize arrangement of the data on the updated archive tape and the tape's characteristics are tabulated below.

Density	= 6250 BPI
Track	= 9 Track
Code	= ASCII
Label	= No label
File	= 1
Blocksize	= 800 characters
Logical Record Size	= 80 characters

Cards are arranged on tape sequentially so that card #1 is the Columbia, Missouri, surface observed weather data and cards #2, #3, and #4 are the WSP sky cover data in eighths for 185 sectors.

7.2.2 Processing of WSP Data for Sky Dome Cloud Correlation Structure

The processing flow of WSP data for finding sky dome cloud correlation structure, Section 7.2.3, is also depicted in Figure 7.1 and is described next.

7.2.2.1 Program CM

By interacting, INTER, with a PC, program CM is called upon to read the WSP data file CMOUT and to compute the correlation of sky cover in each WSP sector for the four seasons--winter (1), spring (2), summer (3), and fall (4). The seasonal results are stored in four files names WS11, WS21, WS31, and WS41 meaning whole sky correlation values for season n version 1. These files are then copied over to the VAX computer for data plotting capabilities on a laser plotter.

7.2.2.2 Program PLOTW

Program PLOTW is a FORTRAN program written on a VAX computer that is called upon interactively to produce plots of cloud correlation structure over a sky dome. The data used to create the plots are the correlation coefficient values stored in the seasonal WS files described in Section 7.2.2.1. The program utilizes the NCAR contour plotting routine CONRAN to

contour the correlation found in each whole sky sector. Description and results of the correlation plots are described in the next section.

7.2.3 Correlation of Sky Cover Over a Sky Dome

7.2.3.1 Introduction

Information on the correlation structure of sky cover is important to the development and implementation of realistic cloud/sky cover simulators. One useful method for providing information about the correlation of sky cover over distances is by correlating ground observed sky conditions at a given anchor station with observed sky conditions at other stations distances apart. This station-to-station spatial correlation is useful in defining the structure of sky cover over large areas but does not address the problem of determining sky cover structure within a single sky dome. This section discusses the application of whole sky photos (WSPs) to the problem of determining realistic spatial correlation of sky cover within a sky dome.

7.2.3.2 Procedures

The procedure for determining spatial correlation of sky cover within a sky dome starts with defining a dichotomous sky condition for each sector of a WSP to be either 1) clear to

partly cloudy (0-4/8) or 2) mostly cloudy to cloudy (greater than 4/8). Then, given the cloud amount observed in the center sector, tabulate, using 2 x 2 contingency tables, the condition of the sky in each of the surrounding sectors. At the completion of the conditional tabulations for each season, compute the tetrachoric correlation for each sector, contour and display the results over a grid showing the approximate center position of each sector.

7.2.3.3 Results

The structure of the resulting correlation of sky cover within the sky dome over Columbus, Missouri, for spring, summer, fall, and winter is portrayed in Figure 7.2. Correlation over the sky dome in general is very high (greater than .84 in the summer and .93 in the winter). Also, correlation decrease over the sky dome appears nearly isotropic for all four seasons; for example, at Columbia, Missouri, the correlation of overhead cloud conditions with conditions in areas off zenith is independent of the azimuth angle. The summer season reflects a steeper rate of decrease in correlation that is most likely caused by the predominant presence of summer fragmented cloud types such as cumulus humulus or stratocumulus. The flatter correlation structure shown for winter reflects the presence of predominant stratoform cloud types associated with large storm systems.

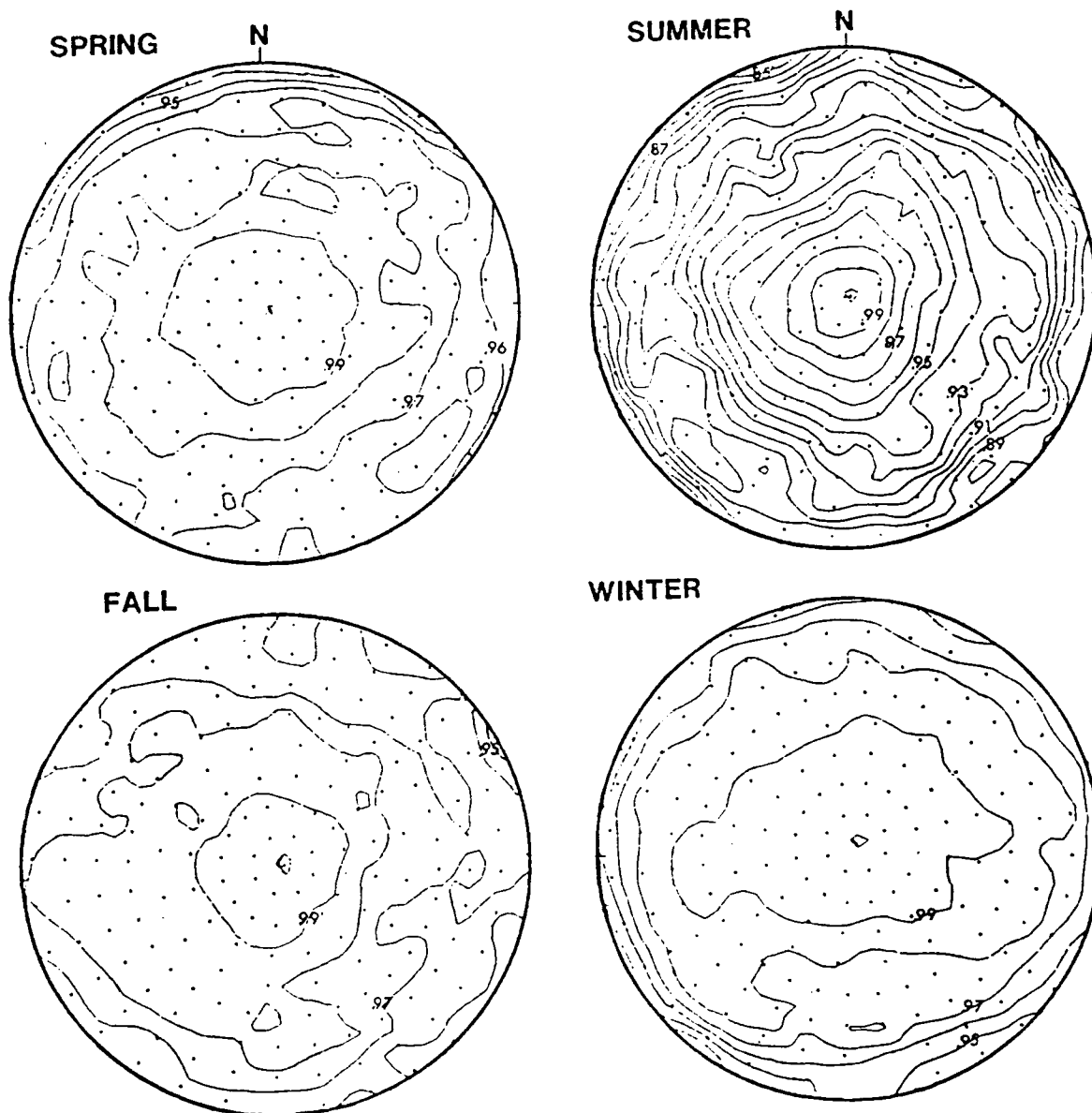


Figure 7.2. Seasonal Spatial Correlations of Sky Cover Calculated from the Center Sector of Whole Sky Photographs Relative to All Other Surrounding Sectors.

According to Lund, Grantham, and Davis (1980), the quality of the whole sky photographs used in producing the WSP digital data ensemble was somewhat degraded in the developing process. It should also be noted that the grid shown in Figure 7.0b would have been more effective if sectors were designed as equal area sectors. In addition, equal area sectors would have made the direct computation of mean correlation a possibility. Calculating the mean correlation has shown initial importance in recent ongoing research used in defining sky cover distributions. (Design of future whole sky imaging systems should consider equal area pixel sky cover identification.)

7.2.3.4 Conclusion

Results include three significant findings: 1) spatial correlation of sky cover within a sky dome is very high (it is above .93 in the winter and above .84 in the summer), 2) correlation drops off more with distance off zenith in the summer than in the winter, and 3) correlation appears nearly isotropic for all seasons.

These results have direct application to the simulation of the probability of cloud-free line-of-sight from single locations on earth to multiple satellites. Any number of satellites over an assigned sky dome area can be evaluated on the basis of the probability of the success or failure of

being seen or communicated with by the appropriate
corresponding ground-based operation.

The section, 7.2.3, has been submitted for publication in
the July 1991 issue of The Journal of Applied Meteorology.

8.0 GEOGRAPHY DATABASES

In the development of an automated global cloud climatology, a background map of outlined major coastal boundaries and/or global elevation data is of value for display of selected cloud parameters globally and for incorporation of terrain induced cloud effects. This section describes two sources that are candidates for mapping major coastal boundaries on a small PC. The first is the modified NCAR vector geography database, and the second is the NAVY 10-minute elevation data (Joseph, 1982) which is in raster format.

8.1 A Modified NCAR Vector Geography Database for Small PCs

A modification to a NCAR geography database has been made so that coastal boundaries of major land areas can be plotted on a small PC. The database used here is one of the earlier versions of NCAR's (Wright, 1978) world geography databases that was digitized at about every half degree resolution. The digitized values were converted to latitude and longitude values representing major continental boundaries, islands, and outlines of the United States boundaries.

The database in its original form is about 300,000 bytes in size. Latitude and longitude points in hundredths of

degrees are stored in variable length blocks as ASCII character coded format. Each block contained a header record describing the latitude-longitude boundaries of the vector within the block. These headers were removed. This reduced the data set down to approximately 240,000 character coded bytes. Ninety degrees have been added to all latitudes. Thus, latitudes range from 0 degrees at south pole to 180 degrees at the north pole. Thus, a latitude value of -85.32 degrees is added to 90, multiplied by 100, and stored as an integer value of 468.

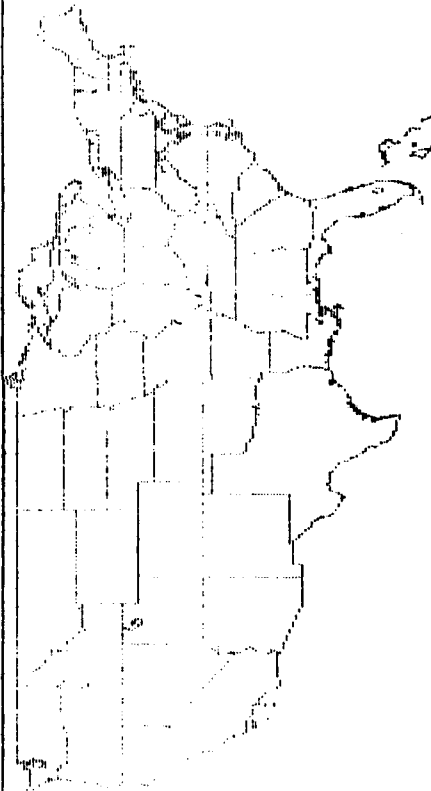
The first vector in the database is the coast of Antarctica. Each block of vectors ends with 400 added to the latitude value in order to signal programs using the data that this is the last point in the block and to go to the next point in sequence with a pen up or beam off.

A sample output of coast lines plotted between 80°S to 80°N is shown in the cylindrical equal area map projection in Figure 8.0a. A separate file containing only the United States area is plotted as shown in Figure 8.0b.

Both the world geography data and the U.S. sector are stored on a single standard 5.25-inch floppy disk. They are stored either as 16-bit unsigned binary integers or ASCII



(a)



(b)

Figure 8.0(a). Cylindrical Equal Area Projection (80°S - 80°N), from the World.DAT Data Set and (b) U.S. State Boundaries, from the U.S.DAT Data Set.

formatted type. The data set names, number of bytes, and type of storage are listed below.

<u>Name</u>	<u>Bytes</u>	<u>Storage Type</u>
US.DAT	130,476	ASCII
US.MAP	43,492	BINARY 16-BIT (21,746 words)
WORLD.DAT	98,928	ASCII
WORLD.MAP	32,976	BINARY 16-BIT (16,488 words)

8.2 NAVY Global 10-Minute Raster Elevation Data

The NAVY 10-minute elevation database contains nine cartiographic parameters of the earth's surface (Table 8.0). These parameters are extracted at a 10-minute x 10-minute resolution and are stored on magnetic tape packed in 30 x 30 cell sizes within 5 degrees latitude-longtitude grid boxes.

TABLE 8.0
Topographic Parameters Within the Original
NAVY 10-Minute Data

Parameter	Description
1	Estimated number of significant ridges
2	General direction of ridges
3	Terrain elevation-modal height
4	Terrain elevation-maximum height
5	Terrain elevation-minimum height
6	Primary characteristics of terrain
7	Secondary characteristics of terrain
8	Percent water surface
9	Percent urban development

These parameters will be useful in defining global elevations in future automated cloud climatologies.

Detailed documentation of the data on magnetic tape can be found in Joseph (1982).

8.2.1 Land and Water 10-Minute Data Set

During a study by Schaaf and Ward, et al. (1990), the terrain height parameters from the Navy 10-minute elevation data set were extracted and put into a format useful for

displaying raster maps of land and water surfaces over the globe. The raster lines are oriented in a south-to-north direction for the southern hemisphere and north-to-south direction for the northern hemisphere. The raster lines are made up of pixels with 10-minute resolution that are coded 0 over water areas and 1 over land areas.

Each hemisphere consists of 6 x 90 degrees or 540 pixels along a raster line oriented by latitude and 6 x 360 degrees or 2,160 raster lines in longitude. Storing each pixel in a byte add up to be over one million bytes (2,160 x 520 pixels per hemisphere). This could not be utilized efficiently on a small PC. Therefore, it was decided to pack the pixels into run-length code in order to compress the data down to a more manageable size.

The original land-water database is divided into two parts, southern and northern hemispheres. A run-length coding procedure is used to compress the data in each hemisphere. The description of compaction begins with the southern hemisphere.

8.2.1.1 Southern Hemisphere Compaction

To save on bits, the compacted database for the southern hemisphere implies that the first run-length value is over

land since the South Polar region is mostly land. Twenty-four land pixels over the Antarctica are observed before reaching the ocean water. Thus, 24 pixels along a raster line of 540 pixels can be ignored. This leaves 516 pixels to worry about in each raster scan. Given this information, the run-length coding scheme begins.

Pixels in a scan run from the South Pole to the equator. Therefore, run-length values run south to north in an oscillating pattern of first land then water then land then water, etc. Programs using this data must keep a cumulative count of the number of pixels being set. When this sum is equal to 540, then the next raster scan in sequence should begin. There are 2,160 scans describing the southern hemisphere. Table 8.1 shows in detail the organization of the actual southern hemisphere run-length coded raster lines 1 and 2. Note that 24 pixels must be added to the first run-length value in each sequence to account for the rest of the land area south of 86 degrees south.

Southern hemisphere latitude can be computed from the cumulative run-length codes using equation 8.1.

$$SHLAT = CRC/6 - 90 \quad (8.1)$$

where

SHLAT = Southern hemisphere latitude

CRC = Cumulative run-length code

Longitude in degrees east can be computed using equation
8.2.

$$\text{ELON} = \text{SL}/6 \quad (8.2)$$

where

ELON = Hemisphere longitude in degrees east

SL = Scan line number (1-2160)

Note that zero has shown up as a run-length code. This is purposely done to keep the land-water sequence intact while keeping the actual values of run-length codes below 256 for single 8-bit byte storage.

The final southern hemisphere data set is made available in two forms on a floppy disk. The first form is an ASCII version that consists of a total of 62,648 bytes in I4 format. The second version is a binary version of 8-bit binary bytes, one byte per value, for a total of 15,669 bytes.

TABLE 8.1
Organization of Run-length Codes for Southern
Hemisphere Scan Lines 1 and 2

#	Run-length Code
1	start scan 1 88+24 land start 90 deg. lat. south; 0 deg. long. east
2	173 water
3	0 1
4	end scan 1 255 w sum=540; stop equator
5	start scan 88+24 1 start 90 deg. lat. south; 6 min. long. east
6	173 w
7	0 1
8	end scan 255 w sum=540; stop equator
	.
	.
	.
	scan 2160 360 deg. east

8.2.1.2 Northern Hemisphere

Since the North Pole consists mostly of water, we start run-length coding over water at 85 degrees north. Thus, 30 pixels can be ignored in the run-length coding scheme, leaving 510 pixels to run-length code along a raster scan that runs from north southward to the equator. Unlike the southern

hemisphere, codes will run as water then land then water then land, etc. Organization of actual northern hemisphere run-length coded raster scan lines 1 and 2 are shown below in Table 8.2.

Thirty pixels of water code must be accounted for at the beginning of each scan in the northern hemisphere.

Given a cumulative run-length coded value, latitude for the northern hemisphere can be computed through use of equation 8.3.

$$\text{NHLAT} = 90 - \text{CRC} / 6 \quad (8.3)$$

where

NHLAT = Northern hemisphere latitude

CRC = Cumulative run-length code

Longitude in degrees east can be computed using equation 8.2

The northern hemisphere run-length coded land-water database is stored on a single floppy disk in two separate formats. The first form is an ASCII version that consists of 109,664 bytes in I4 format and the second is the binary 8-bit byte storage, which totals 27,416 bytes.

TABLE 8.2
Organization of Run-length Codes for Northern
Hemisphere Scan Lines 1 and 2

#	Run-length Code
1	start scan 1 187+30 water start 90 deg. north; 0 deg. long. east
2	19 land
3	6 w
4	59 l
5	5 w
6	3 l
7	14 w
8	184 l
9	stop scan 1 33 w total=540; stop equator
10	start scan 2 189+30 w start 90 deg. north; 6 min. long. east
11	3 l
12	1 w
13	13 l
14	5 w
15	59 l
16	7 w
17	1 l
18	15 w

TABLE 8.2 (Cont'd.)
Organization of Run-length Codes for Northern
Hemisphere Scan Lines 1 and 2

#	Run-length Code
19	183 l
20 stop scan 2	34 w total=540; stop equator
	.
	.
scan 2160	. 360 deg.

A plot of run-length land (black) and water (white) over a section of North America is shown in Figure 8.1.

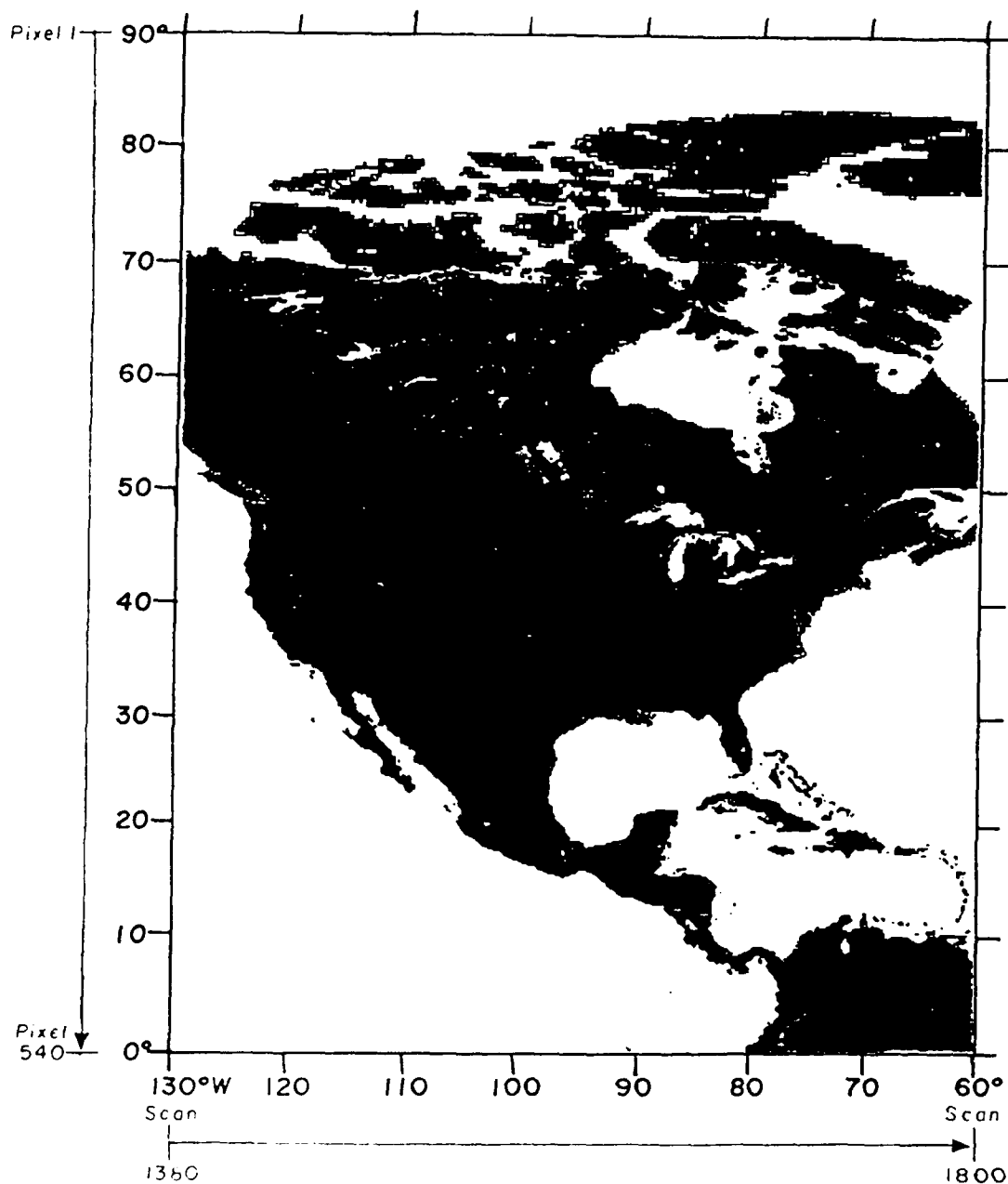


Figure 8.1. Plot of Land-Water Run-Length Code for a Section of North America. The Section Was Extracted from the Northern Hemisphere File for the Area Between 130°W or 230°E to 60°W or 300°E Which Converts to Scan Lines 1380 through 1800. Pixels Run Southward from 90°N to the Equator (1 to 540 Pixels). The Map Projection Shown is Linear in Both Latitude and Longitude.

REFERENCES

Burger, C. F., 1985: World Atlas of Total Sky Cover. AFGL-TR-85-0198, Air Force Geophysics Laboratory, Hanscom Air Force Base, MA, ADA-170474.

Burger, C. F., and I. I. Gringorten, 1984: "Two-Dimensional Modeling for Linear and Areal Probability of Weather Conditions." AFGL-TR-84-0126, Environmental Research Papers No. 875, Air Force Geophysics Laboratory, Hanscom AFB, MA, ADA147970.

Dixon, W. J., chief editor, M. B. Brown, L. Engelman, M. A. Hill, R. I. Jennrich, 1988: BMDP Statistical Software Manual to accompany the 1988 software release. University of California Press, Berkeley, CA.

Epstein, E. S., 1991: "On Obtaining Daily Climatological Values from Monthly Means." J. of Climate, Vol. 4, 365-368, March 1991.

Gerlach, A. M., ed., 1988: Objective Nephology, AFGL-TR-88-0109, Contract F19628-87-C-0046, Systems and Applied Sciences Corporation, ADA-200500.

Gringorten, I. I., 1987: The 3D-BSW Model Applied to Climatology of Small Areas and Lines. AFGL-TR-87-0251, Air Force Geophysics Laboratory, Hanscom AFB, MA ADA-119114.

Hahn, C. J., 1987: Climatological Data for Clouds Over the Globe from Surface Observations. Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO.

Hwang, P. H. et al., 1988: "An Overview of the NIMBUS-7 THIR/TOMS Cloud Data Archive." Submitted for publication to the Bulletin of the American Meteorology Society.

Joseph, D., 1982: Data Format for Global 10-Minute Elevation Data from the U.S. Navy. NCAR, Data Support Section.

Lund, I. A., 1977: Letter Documentation of the Whole Sky Photo Data Set. Air Force Geophysics Laboratory (AFSC), Hanscom Air Force Base, MA 01731.

Lund, I., D. Grantham, and R. Davis, 1980: "Estimating Probabilities of Cloud-Free Fields-of-View from the Earth Through the Atmosphere." J. Applied Meteorology, 19 (4), 452-463.

Lund, I. A., and M. D. Shanklin, 1972: "Photo-grammetrically Determined Cloud-Free Lines-of-Sight Through the Atmosphere." J. Applied Meteorology, 11, 773-782.

Barker-Schaaf, C., and J. M. Ward, et al., 1990: The Hemispheric Eighth Mesh Terrain Elevation and Geography Data Sets. GL-TR-90-0223 Special Reports, No. 263 (ASD), Geophysics Laboratory, Hanscom AFB, MA 01731-5000 ADA-231592.

"Reference Manual," 1977: USAFETAC DATSAV Data Base Handbook. USAFETAC-TN-77-2. Air Force Environmental Technical Applications Center, Asheville, NC.

"Reference Manual," 1986: DATSAV2 Surface USAFETAC Climatic Database Users Handbook No. 4. USAFETAC/UH-86/004, USAF Environmental Technical Applications Center, Asheville, NC.

Smyth, A., J. Willand, and J. Steeves, 1991: Graphical Analysis of Bertoni's LOS Data. Prepared by ST Systems Corporation for Contract No. F19628-88-C-0089, Geophysics Directorate, Atmospheric Structures Branch, Hanscom AFB, MA.

Warren, S. G., C. J. Hahn, J. London, R. M. Chevin, R. L. Jenne, 1986: Global Distribution of Total Cloud Cover and Cloud Type Amounts over Land. Prepared for United States Dept. of Energy, Office of Energy Research, Office of Basic Energy Sciences, Carbon Dioxide Research Division and National Center for Atmospheric Research, Boulder, CO. Technical Note NCAR/TN-273+STR (also DOE/ER/60085-H1).

Wellenger, C. G., 1986: NIMBUS 7 Observation Processing System (NOPS) Tape Specification #T399121 CMATRIX Tape. Prepared by SASC Technologies for the National Aeronautics and Space Administration Goddard Space Flight Center, Greenbelt, MD.

Willand, J. H., 1988: Users Manual for C Cloud S Database. AFGL-TR-88-0060, Contract F19628-87-C-0046, Systems and Applied Sciences Corporation ADA-196497.

Willand, J. H. and A. R. Boehm, 1988: "C Clouds S Climatology of Cloud Statistics." Proceedings of the Cloud Impacts on DOD Operations and Systems, 1988 Workshop (CIDOS-88).

Wright, T., ed. 1978: "NCAR Graphics Software." NCAR Technical Note Fifth edition. National Center for Atmospheric Research, Boulder, CO.

A P P E N D I X A

```

C      FUNCTION TETRA(A,B,C,D)
C      *****
C      ORIGINAL TETRA FUNCTION WAS WRITTEN IN BASIC
C      BY A. BOEHM OF ST SYSTEMS. THIS FORTRAN VERSION
C      WAS ADAPTED FOR THE CDC CYBER 60 BIT COMPUTER
C      BY J. H. WILLAND, ST SYSTEMS, ON 01/03/89.
C
C      PARAMETERS A, B, C, D MAY BE EITHER NORMALIZED ELEMENTS
C      OR ACTUAL COUNTS.
C      REF. BOEHM 1976: TRANSNORMALIZED REGRESSION PROBABILITY,
C      AWSTR-75-259, P24.
C      *****
C      DATA PI/3.1415926536/
C      -----
C      R=-999.0
C      AD=SQRT(A*D)
C      BC=SQRT(B*C)
C      IF (AD.LE.0.0.OR.BC.LE.0.0)GO TO 99
C      ----- FIRST GUESS -----
C      RSIN=SIN(PI/2.*((AD-BC)/(AD+BC)))
C      R=RSIN
C      XN=A+B+C+D
C      P1=(A+C)/XN
C      HQ=-ENORM(P1)
C      P2=(A+B)/XN
C      QK=-ENORM(P2)
C      AN=A/XN
C      R1=0.
C      D1=P1*P2-AN
C      ----- ITERATIVE SOLUTION -----
C      DO 10 K=1,25
C          DQ=FLNORM(HQ,QK,R)-AN
C          IF (ABS(DQ).LT.1.0E-7)GO TO 99
C          RS=R
C          IF (ABS(DQ-D1).LT.1.0E-7)GOTO 99
C          R=R-DQ*(R1-R)/(D1-DQ)
C          D1=DQ
C          R1=RS
C          IF (ABS(R).GE.1.0)R=(SIGN(R,R)+RS)/2.
C      10 CONTINUE
C      99 TETRA=R
C      RETURN
C      END

```

```

C      FUNCTION TFUNC2(A,B,C)
C      *****
C      AVOIDS DIVISION BY ZERO.
C      *****
C      IF (ABS(C).LT.1.0E-8) THEN
C          Y=SIGN(B,B)*(1.0-PNORM(ABS(A)))/2.
C          TFUNC2=Y
C      ELSE
C          TFUNC2=TFUNC(A,B/C)
C      ENDIF
C      RETURN
C      END

```

```

FUNCTION TFUNC(HQ,A)
DATA PI/3.1415926536/
C *****
C BINORMAL SECTOR INTEGRAL.
C REF. YAMAUTI, (1972): STAT. TABLES AND FORMULAS WITH COMPUTER
C APPLICATIONS, JAP STAT. ASSC., PROGRAM 15.
C *****
IF (ABS(HQ).LE.1.0E-50) GO TO 20
A1=ABS(A)
H1=HQ
IF (A1.LT.1.0) GO TO 3
H1=A1*HQ
A1=1./A1
4 AA=A1*A1
HH=H1*H1/2.
D1H=ALOG(HH)
SUM2=EXP(-HH)
SUM1=1.-SUM2
FACT=0.0
S=-AA
DO 10 J=1,80
X=J
FACT=FACT-ALOG(X)+D1H
SUM2=SUM2+EXP(-HH+FACT)
C=S*(1.-SUM2)/(2.*X+1.)
SUM1=SUM1+C
IF (ABS(C).LT.1.0E-10) GO TO 11
S=-S*AA
10 CONTINUE
PRINT 100,HQ,A
100 FORMAT(" TFUNC HQ A",2F10.4,1X,"DID NOT CONVERGE.")
11 T=(ATAN(A1)-SUM1*A1)/(2.*PI)
IF (ABS(A).LE.1.0) GO TO 19
AA=PNORM(-HQ)
HH=PNORM(-H1)
T=(AA+HH)*.5-AA*HH-T
19 IF (A.LT.0.0) T=-T
TFUNC=T
GO TO 99
20 TFUNC=ATAN(A)/(2.*PI)
99 RETURN
END

FUNCTION FLNORM(U,V,R)
C *****
C REF OWEN, 1980: A TABLE OF NORMAL INTEGRALS,
C COMMON STATISTICS.-SIMULA-COMPUTA.,
C B9(4), 389-419, EQ 3.1.
C *****
DATA PI/3.1415926536/
IF (ABS(U).LT.1.0E-8 .AND. ABS(V).LT.1.0E-8) GO TO 999
H=-U
XK=-V
Y=(PNORM(H)+PNORM(XK))/2.
C IF ((H.LT.0. .AND. XK.GE.0. ) .OR. (XK.LT.0. .AND. H.GE.0. )) Y=Y-.5
IF (H.LT.0. .AND. XK.GE.0. ) Y=Y-.5
IF (XK.LT.0. .AND. H.GE.0. ) Y=Y-.5
Y=Y-TFUNC2(H,XK-R*H,H*SQRT(1.-R*R))-
X TFUNC2(XK,H-R*XK,XK*SQRT(1.-R*R))
FLNORM=Y
RETURN
999 FLNORM=.25+ASIN(R)/(2.*PI)
RETURN
END

```

```

      FUNCTION ENORM(P)
      C *****
      C PROBABILITY (P) TO EQUIVALENT NORMAL DEViate END.
      C ALGORITHM AS111, APP STAT 26 1977 P118.
      C INVERSE OF STANDARD NORMAL DISTRIBUTION
      C INTEGRATION FOUND IN PNORM.
      C *****
      Q=P-.5
      IF (ABS(Q).GT.0.42) GOTO 10
      R=Q*Q
      R=Q*( ((-25.44106*R+41.3912)*R-18.615)*R+2.506628)/
X   (((3.130829*R-21.06224)*R+23.08337)*R-8.473511)*R+1.)
      ENORM=R
      RETURN
10   IF (Q.LE.0.0) THEN
      R=P
    ELSE
      R=1.-P
    ENDIF
      IF (R.LT.1.0E-500) R=1.0E-500
      R=SQRT(-ALOG(R))
      R=((2.321213*R+4.850141)*R-2.297965)*R-2.787189)/
X   ((1.637068*R+3.543889)*R+1.)
      IF (Q.LT.0.0) THEN
      ENORM=-R
    ELSE
      ENORM=R
    ENDIF
      RETURN
      END

```

```

      FUNCTION PNORM(E)
      C *****
      C EQUIVALENT NORMAL DEViate TO PROBABILITY.
      C CUMULATIVE NORMAL.
      C INTEGRAL OF THE STANDARD NORMAL DISTRIBUTION
      C FROM INFINITY TO E.
      C *****
      A=ABS(E)
      A=(((((5.621326E-6*A+5.105517E-5)*A+3.968616E-5)*A+3.422739E-3)
X *A+2.2077E-2)*A+5.207516E-2)*A+1.+4.427378E-2)
      A=A**(-16)
      IF (E.GT.0.0) A=1.-A
      PNORM=A
      RETURN
      END

```

```

C      FUNCTION CUB(C,P0,ROBAR,DOF)
C      *****
C      CDC 60 BIT VERSION
C      CUB COVERAGE USING THE BOEHM DISTRIBUTION. 12 JUL 1990.
C      CUB GIVES THE PROBABILITY OF COVERAGE, I.E., THE FRACTION OF A
C      DOMAIN (LINE, AREA, OR VOLUME) THAT HAS THE OCCURANCE OF A WEATHER
C      EVENT. E.G., CLOUDY, BELOW FREEZING, RAIN, ETC.
C      INPUT
C      C=COVERAGE I.E., THE FRACTION OF THE DOMAIN. 0 <= C <=1.
C      P0=MEAN PROBABILITY FOR A POINT IN THE DOMAIN. 0 < P0 <1.
C      ROBAR=MEAN CORRELATION BETWEEN POINTS IN THE DOMAIN. 0<=ROBAR<=1
C      DOF IS THE DEGREE OF FREEDOM RELATED TO THE NUMBER OF PIXELS
C      USED TO OBSERVE THE DOMAIN. 0 <DOF <INFINITY.
C      *****
C      DATA PI/3.1415926536/
C      -----
C      FUNCTIONS USED   ENORM, PNORM
C      -----
C      EM = MEAN OF DOF INDEPENDENT PIXELS
C      FOD=1./DOF
C      EM=ENORM(.5+(C-.5)*(2.**(-FOD)-1.))
C      V IS VARIANCE OF INDEPENDENT CELLS. EVX IS FOR ALL (OR NONE)
C      COVERED
C      EVM IS   FOR HALF COVERED.
C      ELSE INTERPOLATE BETWEEN EVX AND EVM.
C      XMED IS MEAN FOR ALL COVERED.
C      XMED=ENORM(2.**(-FOD))
C      FITSD=P0
C      FITDOF=FITSD**DOF
C      CHECK FOR FITDOF NEAR .5 TO PREVENT DIV. BY ZERO.
C      IF (ABS(FITDOF-.5).GE..001)GO TO 6
C      FITSD=P0*.999
C      FITDOF=FITSD**DOF
C      EVX=((ENORM(FITSD)-XMED)/ENORM(FITSD**DOF))**2
C      IF (C.EQ.1.0 .OR. C.EQ.0.0) THEN
C      V=EVX
C      ELSE
C      SINN=SIN(PI/(2.*DOF+4.))
C      EVM=SINN/(1.-SINN)
C      IF (C.EQ..5) THEN
C      V=EVM
C      ELSE
C      V=(C**2-1.)**2**ALOG(DOF)*(EVX-EVM)+EVM
C      ENDIF
C      ENDIF
C      USE EQUICORRELATED FORMULA TO ADD EFFECT OF CORRELATION
C      BETWEEN CELLS.
C      RMEAN=SQRT(1.-ROBAR)*EM
C      RVAR=ROBAR*(1.-ROBAR)*V
C      RETURN PROB OF STANDARDIZED VARIABLE (X=XMEAN)/XSTD DEV.
C      CUB=PNORM(-(ENORM(P0)-RMEAN)/SQRT(RVAR))
C      RETURN
C      END

```

```

C      SUBROUTINE SKYRO(R, RBAR)
C      *****
C      MEAN CORRELATION GIVEN SCALE DIST. BY A. BOEHM AND J. WILLAND
C      ST SYSTEMS 109 MASS.AV, LEXINGTON MA.
C      VERSION 3.0 2/19/91
C      ASSUMES 2424 KM SKY DOME AND WAVELENGTH=256.R THAT WAS USED IN BAA
C      EDGE IS DISTANCE(IN WAVELENGTHS) ALONG EDGE OF SQUARE WITH 2424 KM ARE.
C      R=SCALE DISTANCE (KM)
C      RBAR=RETURN AVERAGE CORRELATION (RBAR)
C
C      *****
C      DATA N/21/
C      EDGE=SQRT(2424.)/(256.*R)
C      EDGE=.19232/R
C      SUM=(21*21)/4 = 110.25
C      SUM=110.25
C      DX=EDGE/(N-1)
C      DO 20 I=1,N-1
C        DO 10 J=0,N-1
C          SIG=SQRT(FLOAT(I)**2+FLOAT(J)**2)*DX
C          D=SIG - INT(SIG)
C          C=1.-3*D +2.*D*D
C          IF(SIG.GT.1.) C = D*C/SIG
C          SUM=SUM + (N-I)*(N-J)*C
C      10 CONTINUE
C      20 CONTINUE
C      RBAR=4.*SUM/N**4
C      RETURN
C      END

```